NYC Geodatabase Summary

Frank Donnelly, Geospatial Data Librarian, Baruch College CUNY

April 16, 2020

Abstract

This document provides a brief summary of the NYC Geodatabase. For more detailed information, please read the tutorials, data dictionary, and metadata that are posted alongside the database at: `https://www.baruch.cuny.edu/confluence/display/geoportal/NYC+Geodatabase`.

Summary

The NYC Geodatabase (nyc_gdb) is a resource created by the GIS Lab at Baruch College for mapping and analyzing city-level features and data in GIS. It comes in two different database formats whose contents are identical. It includes geographic features and statistical areas from the US Census Bureau, transit and public facility features from various New York City agencies, and census data tables at the PUMA, ZIP / ZCTA, and census tract levels. All features share a common coordinate system: NAD 83 NY State Plane Long Island (feet).

Each database table has a prefix that indicates what type of information it contains, and a column that's designated as a unique identifier or primary key that can be used for relating tables.

- **a tables** contain geographic features; these spatial tables include geometry columns that hold coordinate data. These features are intended for both reference and thematic mapping.
- **b tables** contain "regular" tabular data from the census, and are named for the census dataset, year, and geography. Separate **lookup** *b tables* relate variable column IDs used in the data tables with the actual variable names. *b tables* can be joined to their corresponding *a tables* in order to create thematic maps.
- **c tables** contain geographic features representing legal boundaries, intended for reference mapping only.
- **z table** contains metadata for each table, describing contents and sources.

The databases can be accessed using several different software packages.

- **Open source GIS users** should use the Spatialite / SQLite (.sqlite) format, and access the database with QGIS or the Spatialite GUI.
- **ArcGIS users** should use the personal geodatabase / MS Access (.mdb) format, and access the database with ArcMap or the ArcCatalog.
- **Non-GIS users** can access the tables using the DB Browser for SQLite (.sqlite format) or Microsoft Access (.mdb format). You can export tables out of the database as CSV or Excel files.
- **SQL users** can use either format, but the Spatialite / SQLite format is probably easier for writing and executing both SQL and spatial SQL queries using a number of different interfaces.

QGIS users can connect to the Spatialite database using the browser, and then you can add features from it to the map view. You can also use the Database Manager to connect and view contents of the database, and add features to a project. To map *b table* data, add it and its corresponding *a table* to the project, select the spatial layer, and under its properties select the join tab and create a join based on the matching identifiers. Alternatively you can use the Database Manager to write SQL join statements.

ArcMap users should hit the Add Data Button, and navigate to the mdb database as if it were a folder. Clicking in the database will reveal the individual tables, which you can add to a project. To map *b table* data, add it and its corresponding *a table* to the project, select the spatial layer, and use the joins and relates tab to create a join based on the matching identifiers. Use the ArcCatalog if you want to quickly preview database contents.

Database Contents

Features from the US Census TIGER files

- Boroughs
- Counties (entire metro area)
- Large Public Facilities
- Major Greenspace Areas (parks, cemeteries)
- Roads
- Major Lakes and Coastal Water (latter for entire metro area)

Statistical areas from the US Census TIGER files

- Public Use Microdata Areas (PUMAs)
- Census Tracts (includes NTA codes)
- Population Centroids for Census Tracts
- ZIP Code Tabulation Areas (ZCTAs)

Demographic data from the US Census (updated annually)

- 2010 Census Demographic Profile (for ZCTAs & tracts)
- 5-year American Community Survey data (for PUMAs, ZCTAs, tracts)
- ZIP Code Business Patterns data (for ZCTAs)

Point features from the NYC Facilities Database (updated annually)

- Colleges and Universities
- Libraries
- Hospitals
- Public and Private Schools

Transit point features from MTA / PANYNJ (updated annually)

- PATH Stations (with ridership data)
- Subway Stations and Complexes (with ridership data)
- Train Stations (LIRR and Metro North within NYC only)

Disclaimer: Every effort is made to insure that data published in the NYC Geodatabase, which is compiled from public sources, is processed accurately. The creator, Baruch College, and CUNY disclaim any liability for errors, inaccuracies, or omissions that may be contained therein or for any damages that may arise from the foregoing. Users should independently verify the accuracy of the data for their purposes.