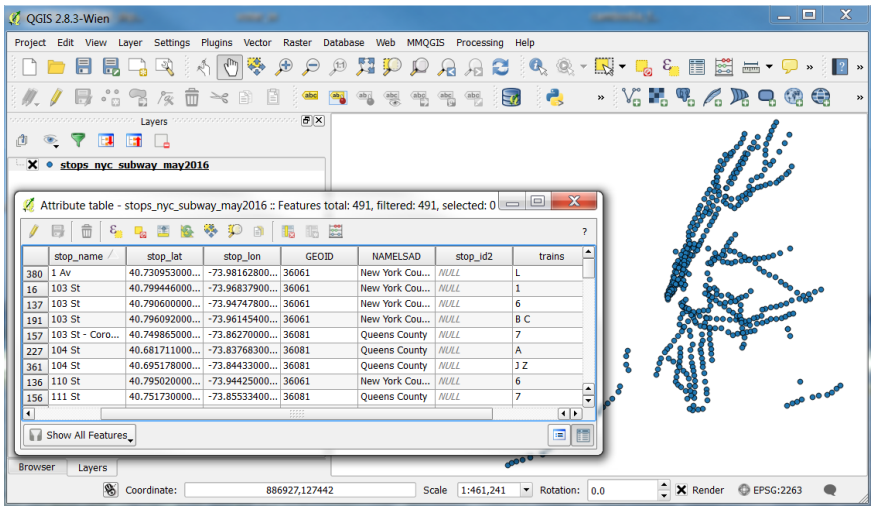
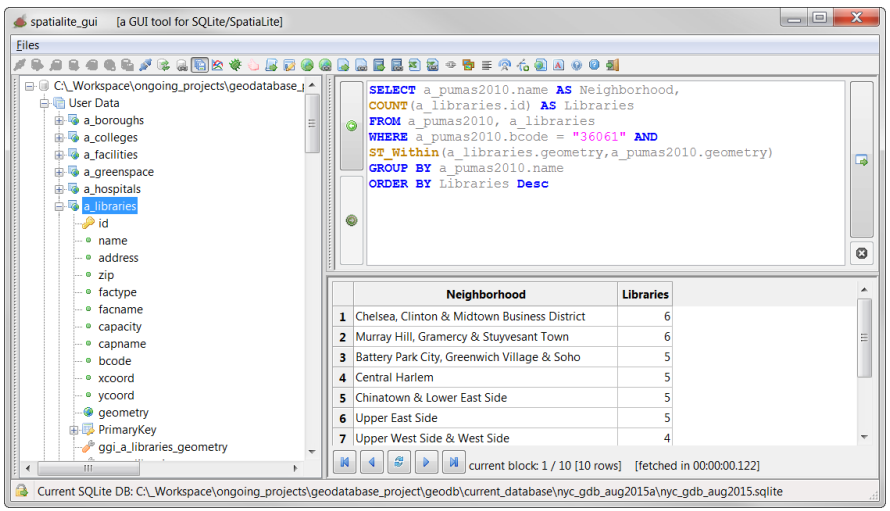
Spatial Data Formats

The Baruch Geoportal houses geospatial datasets that can be visualized and analyzed as map layers in a variety of different software applications. Most of our data is published in either a shapefile, SQLite / Spatialite, or raster format. To learn more about GIS and spatial data, visit our Resources page and the GIS Research Guide for tutorials and workshop / course opportunities.



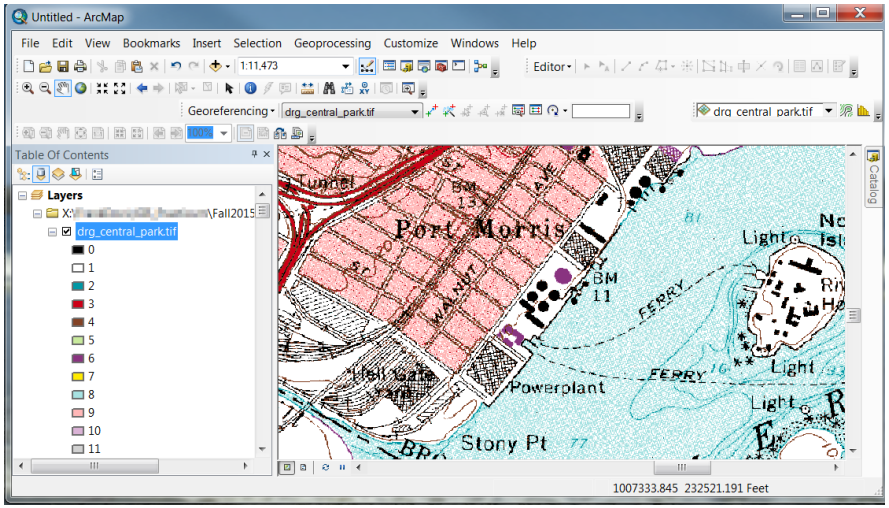
A shapefile and its attributes in QGIS

Shapefiles: files that represent individual geographic features as points, lines, or areas in conjunction with attributes that describe the features. For example, you can have a file for subway stops that contains stop names, trains, and ridership and another file for census areas that contains ID codes and population data. Shapefiles contain geometry and are drawn with a particular coordinate system and map projection that allows them to be depicted visually and overlaid with other geospatial data. Shapefiles can be: used in desktop GIS packages like ArcGIS and QGIS, imported into online web mapping applications like ArcGIS Online, manipulated with scripting languages like Python, imported into statistical packages like R, and imported into databases like PostGIS and Spatialite. Shapefiles contain many individual pieces that perform different functions and need to be kept together. The attribute table contained in the DBF file of the shapefile (without the geometry and spatial components) can be accessed in spreadsheet, database, and statistical programs.



A spatial SQL query in a Spatialite database

Spatialite: this is a file-based database that holds collections of individual point, line, and area features and attributes. Instead of having multiple shapefiles and tables scattered around, all of the data is stored in a single file that can easily be copied and moved. Spatialite is an extension of the SQLite database that has been enhanced to hold and manipulate geometry drawn with a particular coordinate system and map projection. Spatialite databases can be used with desktop GIS packages like ArcGIS and QGIS and can be manipulated with scripting languages like Python. They can also be used as relational or object-oriented databases where you can write SQL queries. There are Spatialite command-line and GUI tools that let you access the full capabilities of a spatial database, but if you just wanted to manipulate the attribute tables (without the geometry / spatial components) you can use a GUI like the DB Browser for SQLite. Data can be exported out of the database as shapefiles or CSV files.



A raster topographic map in ArcGIS

Raster: digital images that consist of maps, air photos, satellite imagery, or gridded data. Rasters are saved in recognizable formats like tifs and jpegs, but are distinct from simple pictures as they are georeferenced: the images have been created or warped to a particular coordinate system and map projection. Raster files can be stand-alone maps or individual pieces of a much larger grid. They can be viewed and used in desktop GIS software like ArcGIS and QGIS as well as in specialized raster software. They can be manipulated with scripting languages like Python and R and some rasters can be imported into statistical packages as grids. If you wanted to view or modify a raster and you weren't interested in the spatial components, you could open it in any graphic design package or picture viewer.