Introduction to ArcGIS Pro II

S4 GIS Institute

Frank Donnelly, GIS & Data Librarian, Brown University

May 16, 2022

Introduction

This tutorial introduces additional, basic features of ArcGIS Pro, building on our previous tutorials which introduced the software interface and covered cartography. We will use our `gis_introduction` project and the Rhode Island tutorial data that we downloaded from the Canvas GIS Institute page. This tutorial will showcase a variety of operations using our theme of studying populations served by public libraries.

Conventions used in this tutorial:

- Summaries of steps appear in **bold face**.
- Names of windows, tabs, and tools appear in *italic face*.
- Names of files, features, and fields appear in `typewriter face`.

Prep Work

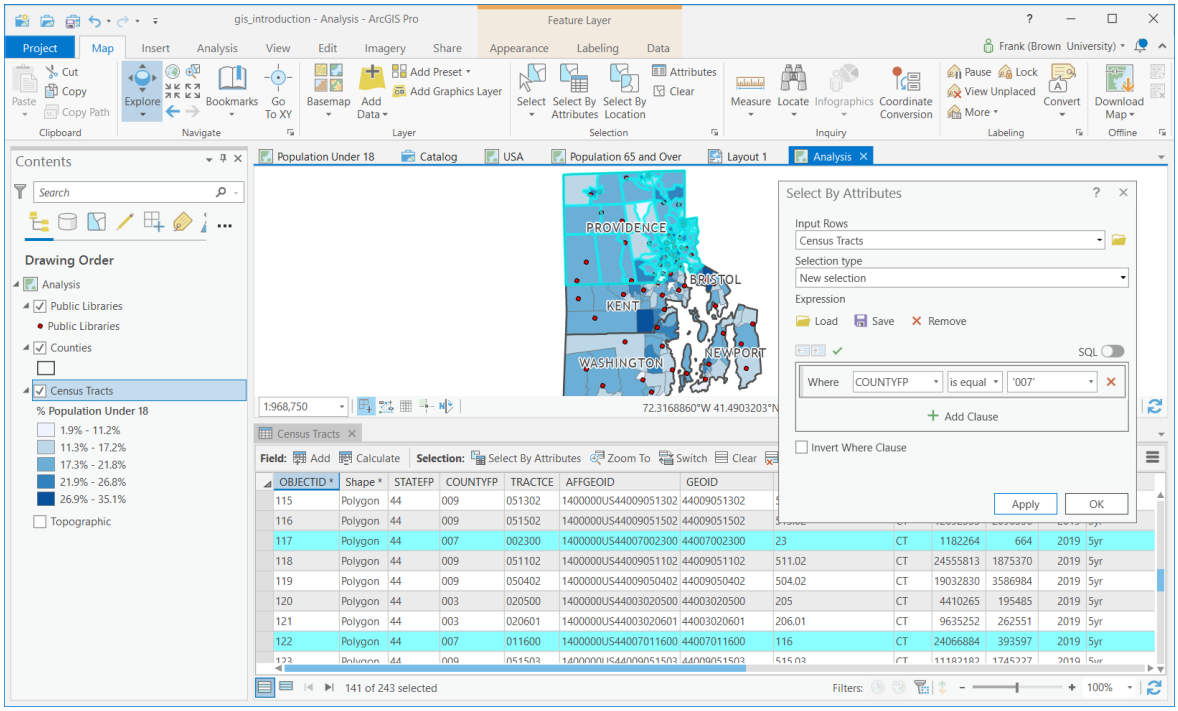
1. **Launch ArcGIS Pro**: And re-open the `gis_introduction` project.
2. **Create a copy of an existing map view**: Go into a *Catalog* view and make a copy of the `Population Under 18` map. Name it `Analysis`. Then open this map view.

1 Tabular Queries and Calculations

The *Select by Attributes* tool allows you to select features based on their attributes, using criteria that you specify in an expression. The features you select will be highlighted, and you have the ability to summarize this information and export the selected features to a new layer.

1. **Preview the attribute table**: Select the `Census Tracts`, right click and open the *Attribute Table*. Examine the table; values saved as text are indented to the left, while quantitative values are right indented.

2. **Select features based on qualitative criteria**: Open the *Select by Attributes* window by either clicking on the button in the *Selection* group on the *Map* ribbon or by clicking the button in the attribute table. Specify `Census Tracts` and the *Input rows*. Keep the *Selection type* as *New selection*, but hit the dropdown to see the various options. Hit the *New Expression* button. Change the column listed in *Where* to `COUNTYFP`, keep *is equal to* and in the last box enter `'007'`. Text fields must be surrounded by single quotes. Hit *Apply*. This selects all the tracts in Providence County, highlighted in blue on the map and in records in the table. The number of selected records appears at the bottom of the map and table views.



3. **Select features based on quantitative criteria**: Change the column to `Under18`, the operator to *is greater than or equal to* and type `1200`. Quantitative values are not quoted. Hit *Apply*. This selects all tracts that have 1,200 or more young people.

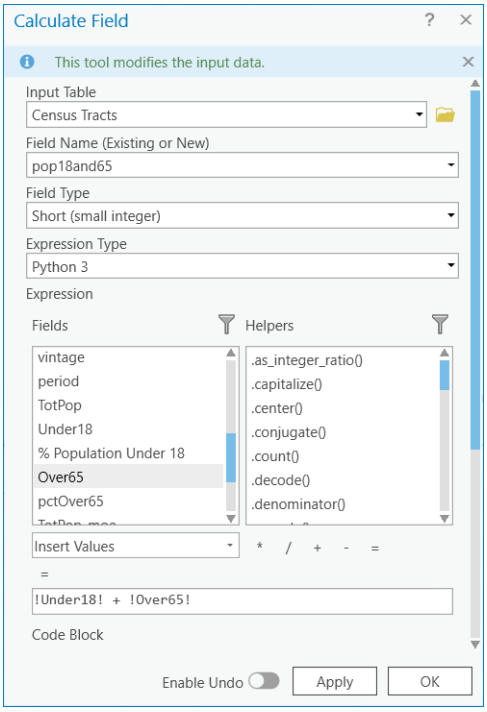
4. **Select features based on multiple criteria**: Hit the *Add Clause* button that appears below the existing clause. In the new clause, specify `COUNTYFP` *is equal to* `'007'`. Hit *Apply*. This selects all tracts that have 1,200 or more young people in Providence County. Note that we used the boolean *And* operator which requires both criteria be met. We could also use the *Or* operator to select areas that meet one criteria or the other. Hit *OK*.

5. **Exporting and clearing**: If we selected `Census Tracts` in the *Contents* pane, right clicked, chose *Data - Export Features*, we could save this file as a new layer, and it will automatically save just the features we selected. We'll skip doing this for now. Hit the *Clear* button either in the *Selection* group on the *Map* ribbon or by clicking the button in the attribute table. This removes the selected features.

2 Table Calculations and Joins

We can edit and modify tables to add new columns and populate them using an expression. Well-formatted data tables stored in a spreadsheet, CSV file, or database table can be added to ArcGIS Pro and joined to GIS layer attributes using a unique identifier they share in common. This allows you to map the data stored in the table, and to integrate its attributes into tabular and spatial queries. We will join a table of Spanish speakers to the tracts layer, as librarians will want to know whether they should purchase Spanish language materials for their collections and offer ESL programming.

1. **Add a calculated field**: With the attribute table of the `Census Tracts` open, hit the *Calculate* button. Keep the *Input Table* as the tracts layer. For *Field Name* we're going to add a new column instead of modifying an existing one. Type pop18and65 in this box. A new *Field Type* option appears below, select *Short (small integer)*. Keep the *Expression Type* as Python 3. In the *Fields* list scroll and click on `Under18` to add this column to the expression list. Hit or type the plus + operator, and then click on the `Over65` field. The new column will contain the sum of these two. Hit the green check button to verify that the expression is correct, and hit *Apply*. Scroll to the right of the table, and you will see the new field.



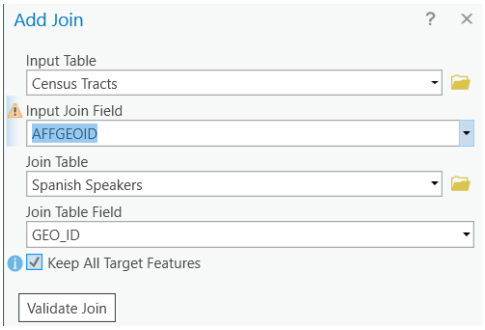
2. **Complex calculations**: Let's calculate a margin of error for this new total. Name the field `pop18_65_moe`, and specify that it is a short integer. In the *Helpers* box click on the option to add *math.sqrt()* to the expression. Click on the appropriate fields and type the appropriate functions and operators to form the expression below, to the take the square root of the sum of squares for the margins of error for the under 18 and over 65 populations:

```
math.sqrt(math.pow(!Under18_moe!,2) + math.pow(!Over65_moe!,2))
```

Hit the green checkmark to verify the expression, then hit *Apply*. Scroll to the right of the table to see the result. Hit *OK* to close the calculator window.

3. **Add spreadsheet to project**: Hit the *Add data* button on the *Map* tab on the ribbon. Browse into the `gis_introduction - intro_sample_data - tables` folder, and select the Excel file named `ri_spanish_2019acs.xlsx`. When prompted, add `Sheet1`. Once the table is added to the *Contents* pane, doubleclick the table and change the name to *Spanish Speakers*. Then open the attribute table and inspect it; it contains the population over 5 years of age that speaks Spanish. It also has a unique identifier that's included in our `Census Tracts` layer.

4. **Join table to features**: Select the `Census Tracts`, right click, and choose *Joins and Relates - Add Join*. The *Input Table* should be `Census Tracts`, and the *Join Table* should be `Spanish Speakers`. Specify `AFFGEOID` as the *Input Join Field* and `GEO_ID` as the *Target Join Field*; these columns contain the same unique IDs. Click *Validate Join* to see if the join will work; there may be a slight mismatch for certain un-populated tracts. Then click *OK*. Open the attribute table for the `Census Tracts`, scroll to the right, and if successful you will see all the fields from the table attached to the attributes of the features.

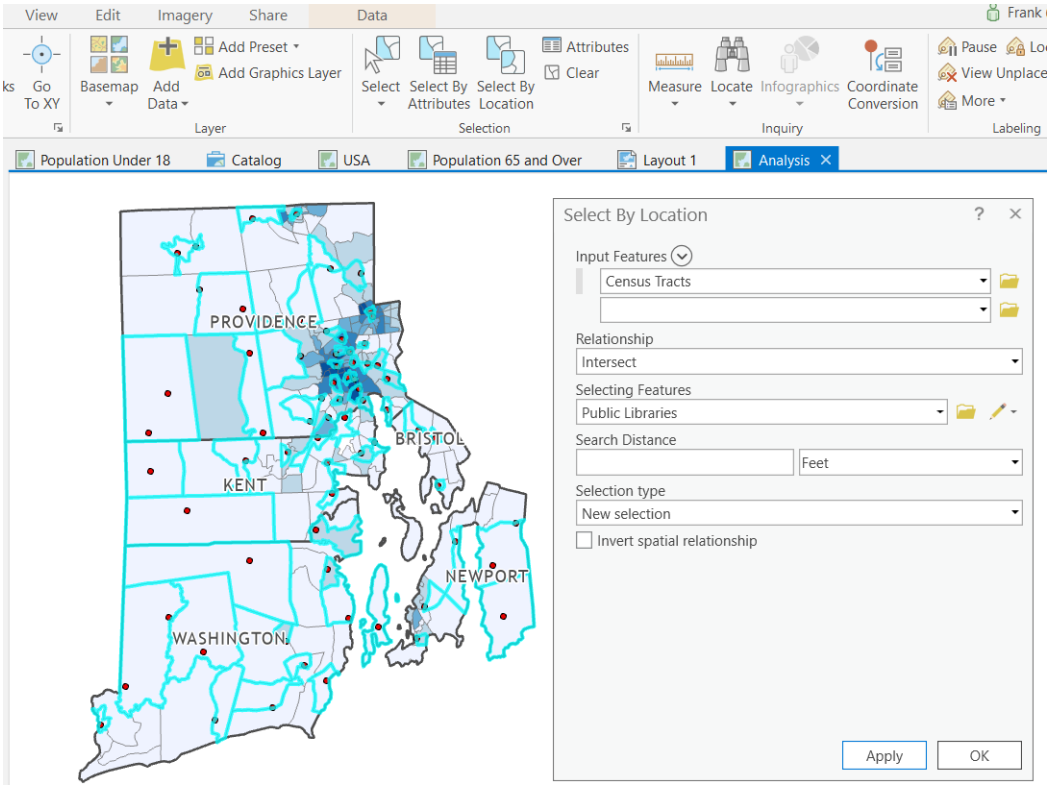


5. **Map joined features**: Open the *Symbology* for the `census_tracts` and map the percentage of the population who speak Spanish, `SPAN_5YPC`. This join exists within this specific project and can be used for mapping, but may not be available for all analytical functions. To make it more permanent, we could import the spreadsheet into the geodatabase as a table and then join it to the features, or we could export the features with the joined table out as a new feature class that will have these attributes fused to it. Joins can be removed by returning to the joins menu for the features.

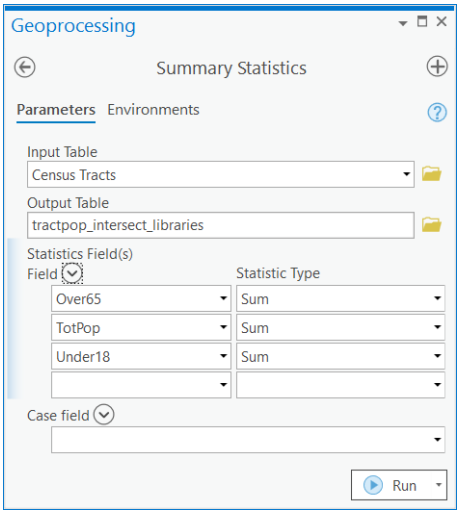
3 Spatial Queries

The *Select by Location* tool allows you to select features in one layer based on their geographic relationship with another layer, i.e. whether features in two layers share a boundary, intersect, or if features from one layer fall inside another layer. The features you select will be highlighted, and you have the ability to summarize this information and export the selected features to a new layer. We will try some different methods for measuring the population that lives close to a library.

1. **Select census tracts with libraries**: On the *Map* tab on the ribbon, click the *Select by Location* button in the *Selection* group. In the *Select by Location* window, select all `Census Tracts` (as the *Input Feature*) that *Intersect* `Public Libraries` (as the *Selecting Features*). Hit *Apply*. Every tract that intersects with a library is selected (in this case, we could also have selected tracts that contain a library and returned the same result). Click *OK*.



2. **Calculate population of selected areas**: Go to the *Analysis* tab on the ribbon, hit the dropdown in the *Tools* group and choose *Summary Statistics*. Specify `Census Tracts` as the *Input Table* and for the *Output Table* store it in the Rhode Island tutorial database and name it `tractpop_intersect_libraries`. Under *Fields* hit the circular dropdown arrow and check: `Over65, TotPop, Under18` and click the *Add* button. Under *Statistic Type* choose *Sum*. Open the summary table to see population counts for the selected tracts.



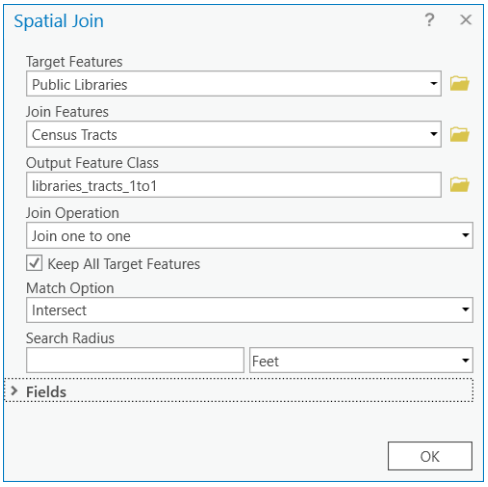
3. **Select census tracts near libraries**: Return to the *Select by Location* window, and now select all `Census Tracts` (as the *Input Feature*) that *Are within a distance of* `Public Libraries` (as the *Selecting Features*). Enter `1` as the *Search Distance* and change the units to *miles*. Hit *Apply*. Every census tract within a mile of a library is selected. Click *OK*.

4. **Calculate population of these selected areas**: Repeat the earlier to step to summarize the population for these tracts. For the *Output Table* store it in the Rhode Island tutorial database and name it `tractpop_1mile_libraries`. Open the summary table to see population counts. This distance method for measuring access captures most of the population, compared to the simple container approach which captured just a fraction. Save your project.

4 Spatial Joins

Spatial joins allow you to take the attributes from one layer and assign them to a second layer based on their geographic relationship. The outcome is a new layer that has attributes of both input layers. For example, we will assign each library the population data of the tract it falls within, or of all nearby tracts.

1. **One to one spatial join**: Select the `Public Libraries` layer, right click, choose *Joins and Relates - Spatial Joins*. Choose: `Public Libraries` as the *Target Features*, `Census Tracts` as the *Join Features*, `libraries_tracts_1to1` as the *Output Feature Class* in the tutorial database, keep the *Join Operation* as *Join one to one*, and the *Match Option* as *Intersect*. This creates a new library layer that has the attributes of the intersecting tract joined to it. Since a library can only fall inside one tract, it's a one to one join. You can optionally choose which fields to take, or take them all. Click *OK*. Open the attribute table of the new layer, and you will see a record for each library with the data for the intersecting tract.



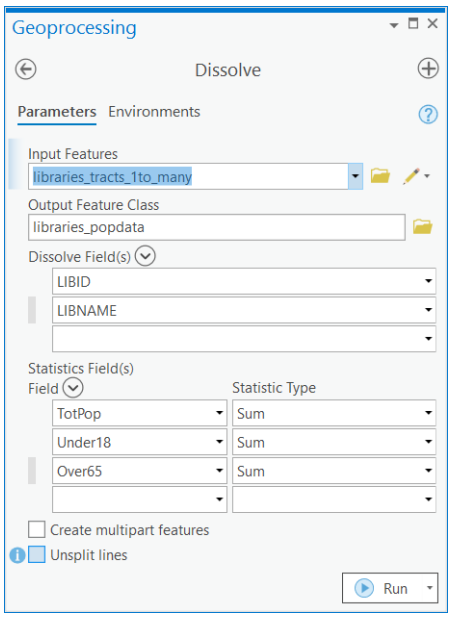
2. **One to many spatial join**: Repeat the last step on the original `Public Libraries` layer, except this time name the *Output Feature Class* `libraries_tracts_1to_many` and modify the *Join Operation* to *Join one to many*, and the *Match Option* to *Within a Distance*. Enter a *Search Radius* of `1 mile`. Click *OK*. This creates a new library layer that has one library record for each census tract that's within one mile of the library; open the attribute table and note that the library attributes are repeated. If you use the *Select* tool and click on a library, multiple library points in this new layer occupy the same space.

5 Geoprocessing

Geoprocessing operations allow you to alter the geometry of a layer, based either on the attributes stored in that layer or on the geographic relationship between two layers, to create a brand new layer. For example, we can dissolve the multiple libraries we created in the last layer and sum their tract population attributes.

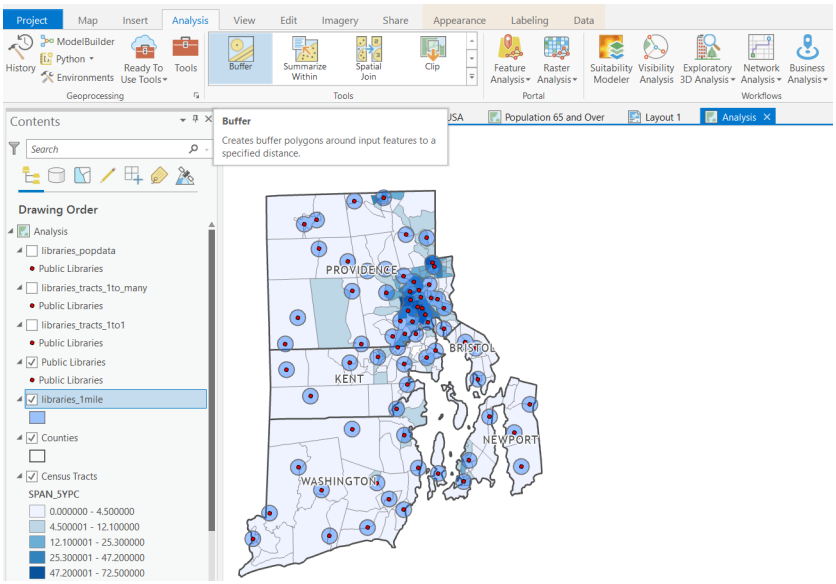
1. **Explore tools**: Click on the *Analysis* tab on the ribbon and click on the *Tools* button. The toolbox contains many different data management and geoprocessing tools; the buttons listed on the toolbar are simply shortcuts to common tools. Drill down through the *Analysis* and *Data Management* toolboxes to explore what's there.

2. **Dissolve libraries to sum population served**: Click on the *Generalization - Dissolve* tool. This allows us to dissolve or aggregate features into larger ones based on shared attributes. For *Input Features* specify `libraries_tracts_1to_many` and for output name it `libraries_popdata`. Choose unique ids that can be used for dissolving the features: `LIBID` and `LIBNAME` are unique to each library outlet, so use those. Under the *Statistics Fields* choose `TotPop, Under18, Over65` and choose *SUM* as the *Statistic Type*. Uncheck the *Create Multipart Features* box (as that won't apply as our output will be points). Click *Run*. Open the attribute table for the output, and you'll see we now have one record for each library, with the data from the tracts summarized so we can measure the potential population served.



3. **Dissolve tracts to sum population for counties**: To better see the power of this tool, do another dissolve. This time, dissolve the `Census Tracts` using the `COUNTYFP` field as the dissolve field, and name the output `county_pop`. Sum the same statistics, but this time check the box to *Create Multipart Features*, as the counties will contain multiple polygons. Click *Run*. You'll see that the actual tract boundaries were dissolved and the new layer represents counties. Open the table and you'll see five records with summed populations (when finished, uncheck this layer in the *Content* pane to turn it off).

4. **Create library buffers**: What would a 1-mile radius around each library look like? On the *Analysis* tab on the ribbon there is a shortcut button to the *Buffer* tool. Specify `Public Libraries` as the *Input Features* and name the *Output Features* `libraries_1mile`. Add a distance of 1 mile, and for *Dissolve Type* specify *Dissolve all output features into a single feature*. Click *Run*. In the *Contents Pane*, make sure the new buffer layer is listed below the *Public Libraries* layer so you can clearly see both. This provides a clear illustration of how tracts within a mile of a library were selected in our earlier example. Save your project.

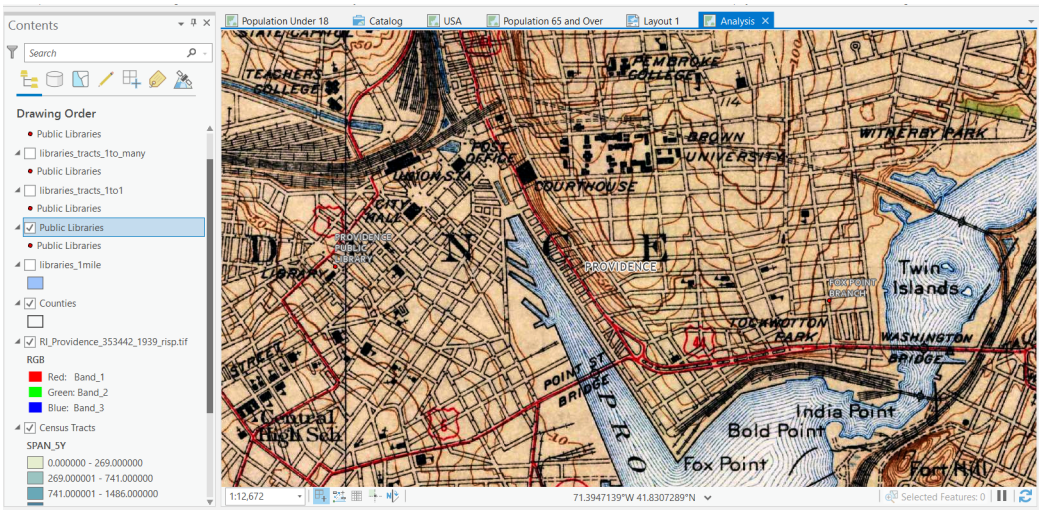


6 Adding Raster Layers

Vector layers are discrete features where strings of coordinates are stored to form points, lines, and polygons, and attributes of these features are stored as records in a table. In contrast, raster layers represent a continuous surface composed of grid cells of equal size, where the cell has a value that denotes some attribute. Some rasters have multiple bands so that multiple values can be stored. The tools for viewing and manipulating rasters are completely different from vectors. Raster analysis is beyond the scope of this tutorial; this section provides a basic illustration.

1. **Add raster layer**: Hit the *Add data* button on the *Map* ribbon, drill down into the `gis_introduction - raster` folder, select the `RI_Providence_353442_1939_risp.tif` file and hit *OK*. In the *Contents* pane drag the raster so that it appears above the `Counties` but below the `Census Tracts`. This is a USGS topographic map of Providence from 1939.

2. **Zoom and explore**: Right click on the raster in the *Contents* pane, and click *Zoom to Layer* to change the map view so it focuses on just this layer. Then, right click on the raster again and choose *Zoom to Source Resolution*. This zooms in to the optimal resolution for this particular raster. If you zoom in even closer, you will begin to see the individual pixels of this layer. Pan around and explore the map. Notice that you can see our `libraries` layer on top of the raster (turn on the labels for the `libraries` if you want to see them more clearly). If you overlay the `tracts` on top of the raster, you can make them semi-transparent to see both at once (select `tracts` layer, *Appearance* tab on the ribbon, *Transparency* in the *Effects* group).



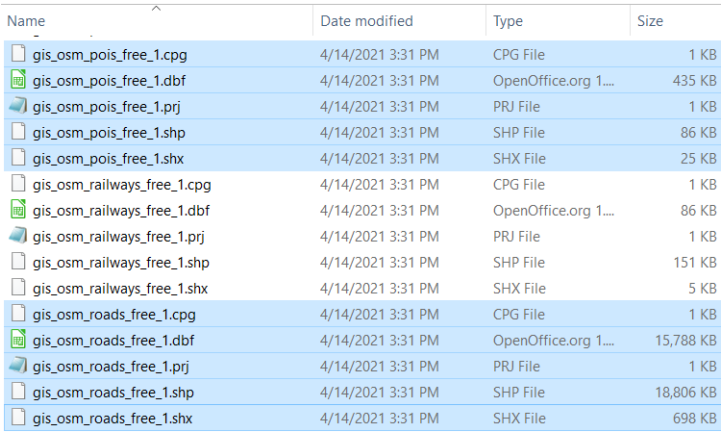
3. **View raster properties**: Right click on the raster and open its *Properties* menu. Under the *Source* menu you can get basic information about the number of cells in rows and columns, their size, and data about different bands. Close the menu when finished. Save your project.

7 Finding Datasets

There is a lot of freely accessible GIS data on the web, produced by government agencies, universities, and non-profits. Datasets are available directly from websites of the data creators and from centralized data repositories. There are also catalogs and directories that list free sources. Not all data is free; there are a number of commercial data vendors and Brown has access to some of these. Visit Brown's GIS research guide for a list of common GIS data sources and access points:

`https://libguides.brown.edu/gis/data`

Vector data is typically provided as individual shapefiles, GeoPackages, and GeoJSON files that contain either a single type of feature or a collection of related features for a given area. Raster data comes in an image file format, and is usually split into a series of tiles that cover a specific area. Vectors and rasters can be added directly to a project, or you can import them into your geodatabase in the *Catalog* view. You may need to re-project files that you download so they match the spatial reference system you are using in your project. Be careful with shapefiles: despite the name they consist of many individual files that must share the same name and exist in the same folder in order to function (see example below). Use the *Catalog* view for renaming and moving files.



Name	Date modified	Type	Size
gis_osm_pois_free_1.cpg	4/14/2021 3:31 PM	CPG File	1 KB
gis_osm_pois_free_1.dbf	4/14/2021 3:31 PM	OpenOffice.org 1....	435 KB
gis_osm_pois_free_1.prj	4/14/2021 3:31 PM	PRJ File	1 KB
gis_osm_pois_free_1.shp	4/14/2021 3:31 PM	SHP File	86 KB
gis_osm_pois_free_1.shx	4/14/2021 3:31 PM	SHX File	25 KB
gis_osm_railways_free_1.cpg	4/14/2021 3:31 PM	CPG File	1 KB
gis_osm_railways_free_1.dbf	4/14/2021 3:31 PM	OpenOffice.org 1....	86 KB
gis_osm_railways_free_1.prj	4/14/2021 3:31 PM	PRJ File	1 KB
gis_osm_railways_free_1.shp	4/14/2021 3:31 PM	SHP File	151 KB
gis_osm_railways_free_1.shx	4/14/2021 3:31 PM	SHX File	5 KB
gis_osm_roads_free_1.cpg	4/14/2021 3:31 PM	CPG File	1 KB
gis_osm_roads_free_1.dbf	4/14/2021 3:31 PM	OpenOffice.org 1....	15,788 KB
gis_osm_roads_free_1.prj	4/14/2021 3:31 PM	PRJ File	1 KB
gis_osm_roads_free_1.shp	4/14/2021 3:31 PM	SHP File	18,806 KB
gis_osm_roads_free_1.shx	4/14/2021 3:31 PM	SHX File	698 KB

Creating vector-based thematic maps is a multi-step process that often involves downloading GIS data files for the geography from one source and attribute data from a separate source that you will associate with these boundary files in a table join. In other cases you may find point-based coordinate data representing features or events stored in a tabular format, and you will need to plot this data in order to map it (*Map* tab on ribbon, *Add Data* dropdown, *Add XY Data*). For resources related to attribute and XY data, visit:

`https://libguides.brown.edu/gis/georesources`

ArcGIS Pro provides direct access to a repository of curated datasets through their *Living Atlas*, as well as an even larger collection of user-contributed data via *ArcGIS Online*. Use the *Add Data* button to search and access datasets in these portals.

8 Experiment on Your Own

1. Run some statistics on the attribute table for the `Census Tracts` to find the mean and median for the percentage of the population that is under 18 and over 65 and the percentage of the population 5 years and over that speaks Spanish. Then run three different attribute queries to select tracts that fall above these means or medians.

2. In the library literature, two-mile buffers are often used for measuring service areas. Create two-mile buffers around the `Public Libraries`, and select all `Census Tracts` that either contain a public library or have their geographic center within the 2-mile library buffer. Then run summary stats on these results to measure the population being served.

3. Look in the *Toolbox* for some distance tools, and calculate the distance from every census tract to the nearest library.