



## CHAPTER 3

# Exploring geospatial relationships

### Exercise objectives

#### 3a: Extract part of a dataset

- Add data to a project.
- Select features by attributes.
- Export the selection to a new dataset.

#### 3b: Incorporate tabular data

- Join data tables.
- Apply informative symbols.
- Import layer symbology.
- Use the Swipe function to compare layers.
- Overlay additional data.

#### 3c: Calculate data statistics

- Add a new field.
- Calculate field values.
- Display a new field.
- Calculate summary statistics.
- Examine infographics.

#### 3d: Connect spatial datasets

- Relate tables.
- Spatially join data.

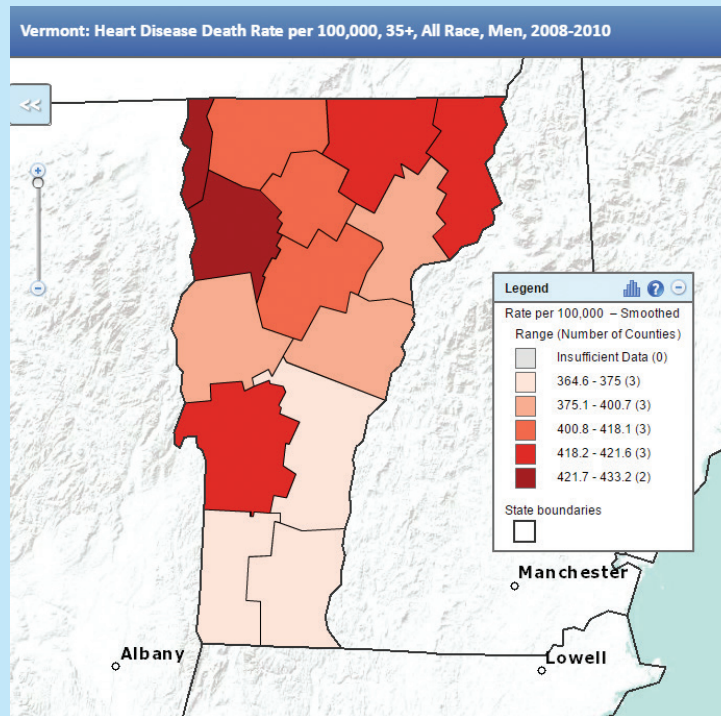
The power of GIS extends far beyond exploring digital maps. You can combine datasets, enrich them with new attributes, derive statistics from them, and obtain new information based on their relationships. In this chapter, you will begin taking advantage of some of the more sophisticated capabilities offered by ArcGIS Pro.

*Scenario:* You have been hired by a state health coalition that is focusing its efforts on raising awareness about and lowering obesity rates in the state of Illinois. Obesity prevalence has risen dramatically throughout the state over the past decade, along with the incidence of type 2 diabetes, but the prevalence of obesity is higher in some areas than others. You will explore obesity prevalence rates by county, create visual aids for displaying a year-over-year rising trend, and begin to explore possible reasons why some counties may have higher prevalence rates than others. For instance, is there a link between average income level and obesity levels? What about comparing obesity rates to data that shows limited access to grocery stores—are these “food deserts,” with lesser accessibility to healthy foods, more likely to have higher obesity rates?

Your job is to combine obesity statistics provided by the Centers for Disease Control and Prevention (CDC) with county GIS data to do an initial visual analysis of the data, which will eventually lead to more rigorous analysis techniques. The end objective of the analysis is to create informative products that will educate and guide advocates and policymakers.

## GIS IN THE WORLD: AN INTERACTIVE ATLAS OF HEART DISEASE AND STROKE INCIDENCE

The Interactive Atlas of Heart Disease and Stroke (<https://nccd.cdc.gov/DHDSPAAtlas>), created by the CDC and powered by GIS, creates a visual guide for anyone who wants to see where cardiovascular disease occurs and offers insight about who is at highest risk. Read more about the project in *ArcNews*, “CDC Aims at Reducing Annual Heart Disease- and Stroke-Related Deaths in the US”: [www.esri.com/about/newsroom/arcnews/cdc-aims-at-reducing-annual-heart-disease-and-stroke-related-deaths-in-the-us](http://www.esri.com/about/newsroom/arcnews/cdc-aims-at-reducing-annual-heart-disease-and-stroke-related-deaths-in-the-us).



**Figure 3.1.** The CDC atlas allows you to apply various filters to get the information you need. Here, the map shows heart disease death rates in Vermont between 2008 and 2010, limited to men above the age of 35. This map was created using the Interactive Atlas of Heart Disease and Stroke, a website developed by the Centers for Disease Control and Prevention, Division for Heart Disease and Stroke Prevention. <http://nccd.cdc.gov/DHDSPAAtlas>.

## DATA

In the C:\GTKAGPro\HealthStudy\Data folder:

- IL\_food\_deserts.shp—a shapefile containing polygons that represent low-income areas with limited access to grocery stores (*Sources: Data and Maps for ArcGIS; USDA*)
- IL\_med\_income.shp—a shapefile representing Illinois counties with an attribute for median income values (*Sources: Data and Maps for ArcGIS; US Census Bureau*)
- Obesity\_Prevalence.dbf—a spreadsheet containing obesity prevalence rates for Illinois counties from 2004 to 2010 (*Source: Centers for Disease Control and Prevention*)
- us\_cnty\_enc.shp—a shapefile containing polygons that represent the county boundaries for the East North Central region of the midwestern United States (*Source: Data and Maps for ArcGIS*)

## Exercise 3a: Extract part of a dataset

*Estimated time to complete: 20 minutes*

In this exercise, you set up your project. You will add data to a new project and then do some basic processing to get your data in shape. Once you have completed these steps, you will be ready for a deeper exploration of your data.

### Exercise workflow

- Open a project and add a shapefile of polygons that represent counties from five states in the East North Central region of the Midwest.
- Select the features with a STATE\_NAME attribute of Illinois.
- Export the selected features to a new dataset that contains counties in Illinois only.

### Add data to a project

1. Start ArcGIS Pro, click Open Another Project, find the C:\GTKAGPro\HealthStudy folder, and open HealthStudy.aprx.

The project opens with a 2D map that includes the Esri topographic basemap centered on North America.

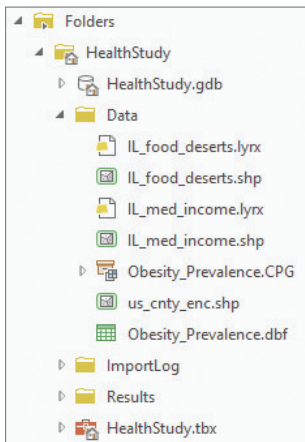
*ArcGIS interface elements change periodically to keep up with evolving software functionality, so the images in this book may not exactly match the software interface you see.*

2. On the Map tab, ensure that the Explore tool is active, and click any area in North America.

Notice that there is no pop-up window. That is because basemaps function the same way as images—they do not have individual features with attributes.

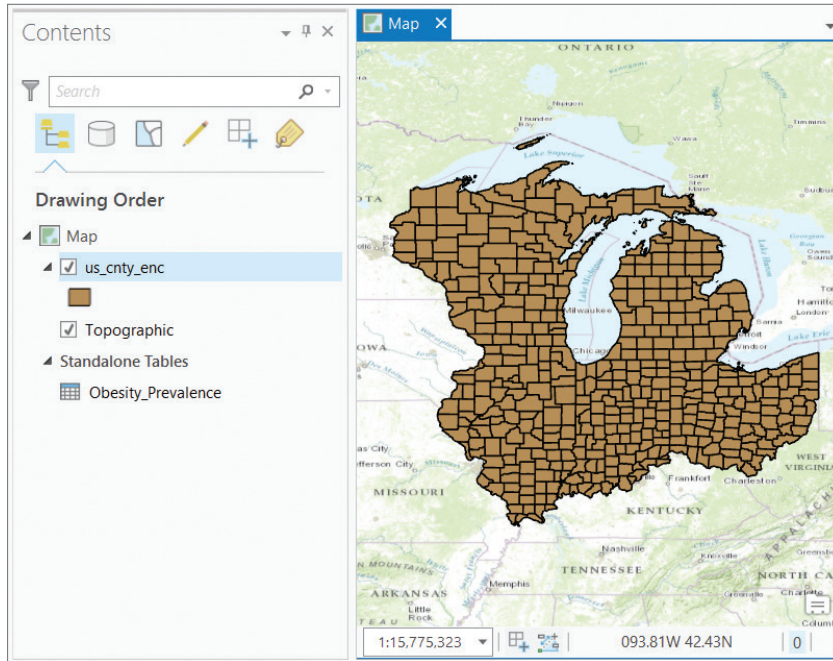
3. In the Catalog pane, expand the HealthStudy\Data folder.

Notice that this folder contains three shapefiles (.shp), two layer files (.lyrx), and one data-base file (.dbf).



4. Press and hold Ctrl, and click us\_cnty\_enc.shp and Obesity\_Prevalence.dbf, and then drag them to the map display.

Your map project now includes a shapefile that contains county features for the East North Central region of the midwestern United States. The area includes five states: Wisconsin, Michigan, Illinois, Indiana, and Ohio. The project also includes a nonspatial .dbf table, listed in the Contents pane under Standalone Tables. Because the table is tabular and does not have spatial attributes, it cannot be displayed on the map in this form.



*The colors will vary because they are randomly assigned.*

- With the `us_cnty_enc.shp` layer selected in the Contents pane, click the Data tab, and then click the Attribute Table button. Examine the table.

---

*What is the field name that indicates the state within which the county features are located? How many residents of Wayne County are between the ages of 22 and 29 years?*

---

- Close the attribute table.

You'll want to limit the `us_cnty_enc.shp` dataset to the state of Illinois only. You can consider a few operations to achieve this subset:

**Definition query.** You can set a definition query to limit the visible counties to only those within the state of Illinois. Definition queries are helpful when you want to work with a subset of data in a map while maintaining the source data. You can set a definition query on the Data tab of the contextual ribbon. However, you must think ahead—what is next in the workflow? You want to append Illinois obesity statistics (which are in a stand-alone table) to a spatial layer of Illinois counties. You will do this by performing a join operation, which is explained in detail in exercise 3b. But with a definition query defined, the source data does not change, so the operation will attempt to join Illinois obesity rates to every county in the five-state dataset. Although this operation works, it takes a little longer, so it is best to go another route.

*You will work with definition queries later in the book.*

**Clip.** You can use the Clip operation to select a portion of `us_cnty_enc.shp` based on another layer—namely, a layer that provides a boundary of the state of Illinois. However, you do not have an Illinois boundary layer now. There is another way.

*You will work with the Clip tool later in the book.*

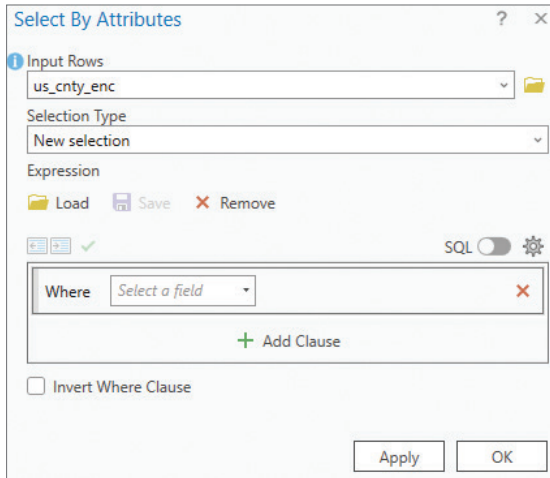
**Select and export.** You can select only those counties that belong to the state of Illinois and export the selected features to a new dataset. Then you will have an exact one-to-one correlation between your spatial attribute table (because the new output will have one record for each Illinois county) and your stand-alone table (obesity statistics). This method is the best for the current study.

## Select features by attributes

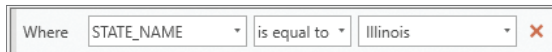
You will create a selection in which the state name attribute value equals Illinois.

1. On the Map tab, open the Select By Attributes tool.  
The Select By Attributes tool opens.

- For Input Rows, click the `us_cnty_enc` shapefile. For Selection Type, maintain the default New selection (click the down arrow to see additional options).



- Click Add Clause, and then choose the correct options to create the following expression: **Where STATE\_NAME is Equal to Illinois.**



1

2

3

4

5

6

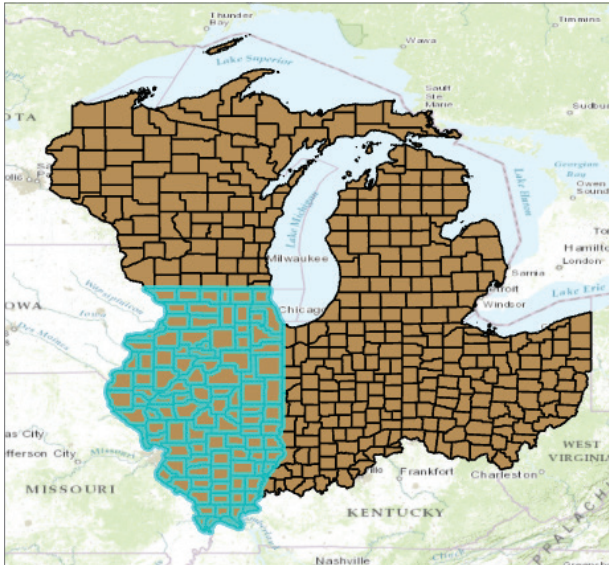
7

8

9

The expression you created is called an **attribute query**—a request for features in a table that meet user-defined criteria.

4. Click the red X to delete any unused fields in the definition query, and click OK. Illinois counties are now highlighted in the selection color (cyan, by default).



Now you will create a feature layer from the selected set of counties.

*You have made a selection based on attributes, but you can also make a selection based on location. For instance, suppose you want to select parcels that are within the boundaries of a flood zone. In this case, use the Select By Location tool. You can choose from many relationship types, including "Completely within," "Have their center in," "Intersect," and "Within a distance." You will use Select By Location later in the book.*

## Export the selection to a new dataset

1. In the Contents pane, select `us_cnty_enc`.
2. On the Data tab, in the Export group, click the Export Features button.

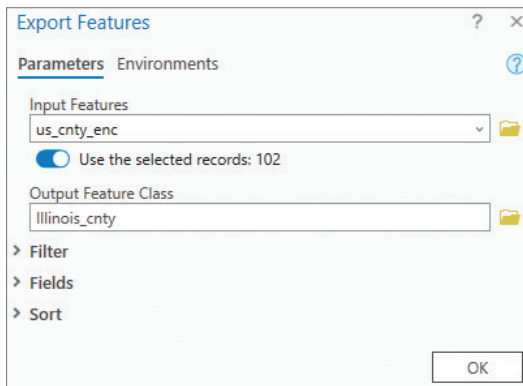
The Export Features tool pane opens, in which you can modify parameters and environments.



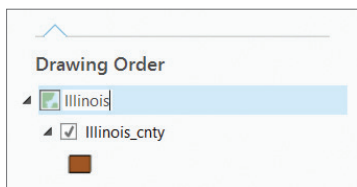
- In the Export Features pane, for Input Features, click `us_cnty_enc`. For Output Feature Class, maintain the default location but change the Output Name to `Illinois_cnty`.

*To see the location of the output feature class, click inside the Output Feature Class box and scroll to the far right. After you have typed the output name, click outside the Output Feature Class box to confirm the output name.*

Note that the new output will be a geodatabase feature class rather than a shapefile. By default, all processing outputs are saved to a project geodatabase—the conversion from shapefile to geodatabase feature class is done behind the scenes.



- Click OK.
- Remove the old shapefile `us_cnty_enc` from the map contents. Zoom to the extent of the `Illinois_cnty` layer. Change the map name to **Illinois**.



### REMINDE ME HOW

- Right-click layer and click Remove.
- Right-click layer and click Zoom To Layer.
- In the Contents pane, click Map once to select it and click again to make it editable. Type **Illinois**, and then press Enter.