Exercise 1

Spatial Data Models

Objective – Explore and Understand Spatial Data Models

1.1 Introduction

In this exercise, you'll explore and manage geospatial data using both the Browser panel and the Data Source Manager of the FOSS4G software QGIS. Both the Browser panel and the Data Source Manager allow you to locate data in your local file system, get some information about the data and and add it to QGIS. As you will see, the Data Source Manager also has a version of the Browser panel available. Therefore, there is a lot of duplicate functionality between the two. Overall the Data Source Manager has more functionality. However, the Browser panel can also be a handy way to add data. It is analogous to Windows Explorer, but works specifically with geospatial datasets.

This exercise will also introduce you to the QGIS interface, which is used throughout the workbook. It is important to learn the concepts in this exercise as future exercises will require the skills covered here.

This exercise includes the following tasks:

- Task 1 – Working with the Browser panel.
- Task 2 – Become familiar with geospatial data models.
- Task 3 – Working with the Data Source Manager.

1.2 Objective: Explore and Understand Geospatial Data Models

Geographic Information Systems model the real world with representations of objects such as lakes, roads and towns. Geospatial data models are the means used to represent these features. They are composed of two parts: spatial features and attributes. When these two components are combined they create a model of reality (Figure 1.1, on the following page).

There are two main geospatial data models: vector and raster. Geospatial features can be represented in either data model.

Vector Data Model – This model is best for representing discrete objects. It comes in three forms: point, line, and polygon. Vector representations have more precision than rasters. This is because points, lines and polygons are defined by X/Y coordinates. Vector is best when precise distances, lengths and areas are needed. It is also ideal for network analyses, for example the shortest distance between two points across a linear network. It is also well suited to cartography since you can use different icons for points, patterns for lines and different outlines and fills for polygons.

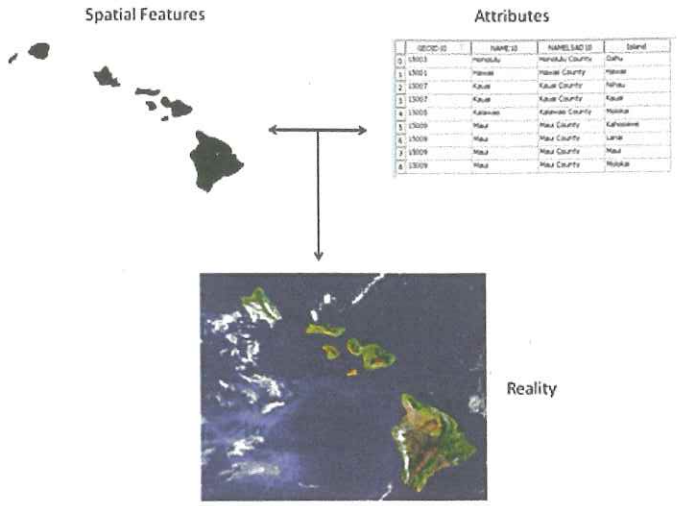


Figure 1.1: Two parts of the geospatial data model

Raster Data Model – This model is ideal for representing continuous phenomena such as elevation or precipitation. A raster is composed of a matrix of contiguous cells, with each cell (pixel) holding a single numeric value. Cells are typically square and have a characteristic known as resolution. This is the dimension of a cell. Typically resolution is expressed in coordinate units such as meters or feet. Rasters are also used for aerial photography and satellite imagery. Furthermore, they are a great choice for site suitability modeling. For example, rasters can be combined via mathematical operations (raster algebra) to compute potential locations for things like species habitat, disease vectors or store locations.

1.3 Task 1 - Working with the Browser panel

In this task, you will become familiar with the Browser panel. The first step in working on a project with geospatial datasets is to organize your workspace. It is important that we organize datasets logically on the computer and make them easy to find. In this task, you will obtain a copy of the exercise data and explore how the data is organized using the Browser panel.

Open QGIS. The way you open QGIS will vary depending on your operating system. For this series of exercises, we will explain how to open and use QGIS using the Microsoft Windows 10 operating system.

1. Click `Start | All Programs | QGIS | QGIS Desktop 3.x`.

> NOTE: You may see more than one version of QGIS. One with GRASS 7 support and one without. GRASS is another open source GIS software package. It has its own data structure named a GRASS database. QGIS with GRASS 7 is configured with support for GRASS databases. Since you are not working with a GRASS database, either version will work.

The interface to QGIS is straightforward. Both the Browser and Layers panels are docked to the left side of QGIS (Figure 1.2, on the next page). Most of the remaining space is taken up by the Map canvas. Above are toolbars and menus.

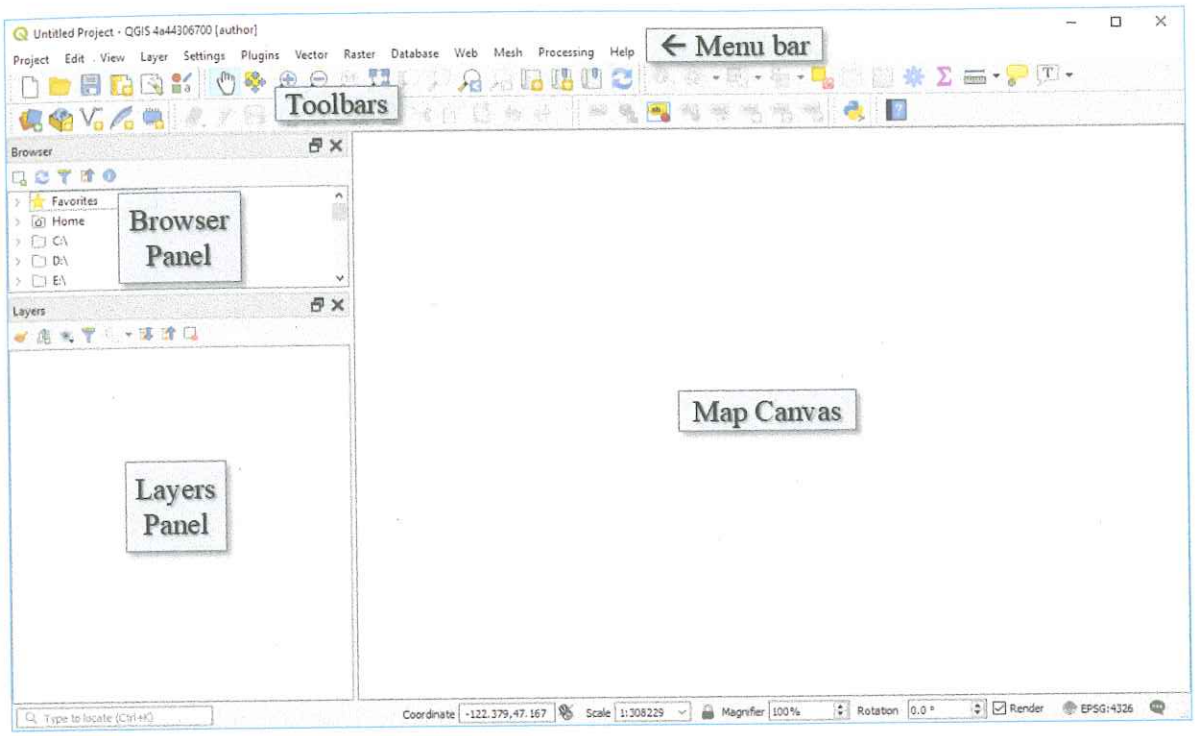


Figure 1.2: QGIS Desktop

Your QGIS window may look slightly different than the one pictured above. To reset your display back to the default settings, click the Settings | Options | System tab | Settings section | Reset button, then click OK and restart QGIS. Also note that many of the QGIS screenshots in this book contain the word [author] in the title bar. QGIS allows you to set up User Profiles. To do so go to the menu bar and choose Settings | User Profiles. Each profile has settings for the plugins installed, toolbars enabled, arrangement of toolbars etc. This book has been written using a Profile named author which has the default QGIS settings. You may also choose to set up profiles for specific analyses, particular projects or clients.

For this task you will focus on the Browser Panel which displays the file tree. It shows your computer's files and folders. Your machine may have a different set and number of drives listed here—this is fine. Below the drives are Database and Web Server connections. There are no connections of either type at this point.

2. Look at the file tree. Click the arrow to the left of the C: drive to expand it. You will now see all of the subfolders directly under the C:/ drive.

3. Expand the `Discover_QGIS_2ndEd\Part_1_Introduction_to_Geospatial_Technology\Exercise_1_Data` folder in the file tree by clicking the arrows to the left of each folder. You will now see the contents of the `Exercise_1_Data` folder for the exercise (Figure 1.3, on the following page).

4. Take a moment to read the names of the files. There are three folders and several files listed with different icons. The vector file icon indicates that the dataset is a vector layer. This icon is used to represent raster data but is also used for other files such as the XML files you see here. The icon indicates that the dataset is a database.

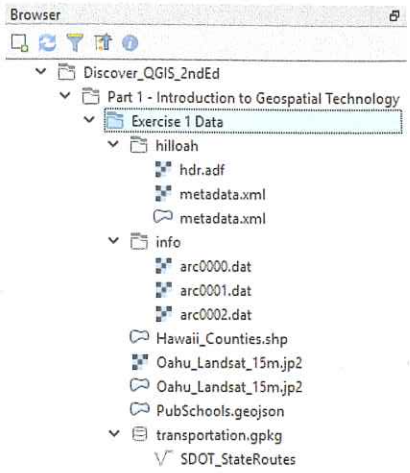


Figure 1.3: Exercise data in Browser panel

Note that QGIS includes different User Interface (UI) Themes. To access these go to the menu bar and choose *Settings | Options* and select the *General* tab. In the *Application* section find *UI Theme*. There is a *Night Mapping* theme which has been greatly improved in recent releases. Using this while working in the evening can make your mapping exerience more pleasant. After switching to one of these you will need to close and relaunch QGIS for the change to take effect. In QGIS 3.6 Noosa there is another new UI them named *Blend of Gray*.

1.4 Task 2 - Become Familiar with Geospatial Data Models

Now that you are familiar with the basic layout of the Browser Panel, you will explore some geospatial data. You will learn how to access properties of a layer and begin adding data to QGIS.

1. Let's take a closer look at these data currently listed in the `Exercise_1_Data` folder.

2. Right-click on the `Hawaii_Counties.shp` layer in the file tree. From the context menu choose Layer Properties (Figure: 1.4). Also notice that there is an option to Export Layer. This allows you to export a layer to a new format without having to first load the layer into QGIS.

Figure 1.4: Browser Panel Context Menu

3. A Layer Properties window opens with some basic information about the dataset. You will notice that the Storage type is ESRI shapefile. You can also see that it has a Geometry type of Polygon(MultiPolygon) and it has 9 features (Figure 1.5, on the facing page).

This Layer Properties window also has a Preview tab for seeing a preview of the spatial features and an Attributes tab for previewing the attribute table.

4. Close the Layer Properties window.

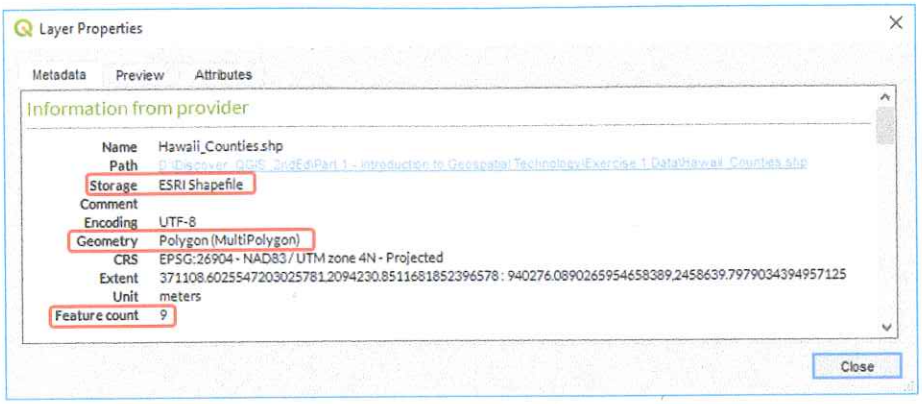


Figure 1.5: Layer Properties

In addition to data models (vector and raster) we have to understand file formats. Some file formats are designed to store vector, and others raster data. Shapefiles are vector file format. In fact, they are probably the most common vector file format. An individual shapefile can only contain one geometry type (polygon, line, or point). A shapefile is actually a collection of files on the computer with a common name, but different extensions. There are three files that are mandatory to have a functioning shapefile:

- .shp -- stores the feature geometry
- .shx -- an index linking the .shp to the .dbf
- .dbf -- a database file which stores the attributes

The shapefile format is considered a mostly open specification. The specification is published but was not developed in an open forum. The format, while ubiquitous, has some significant limitations: A) Attribute column names are limited to 10 characters, B) lacks a time data type, C) only supports text fields of 255 characters in length, D) is limited to 2 GB in size and E) there is a maximum number of attribute columns of 255. The Esri whitepaper describing the format can be found here: `http://esri.com/library/whitepapers/pdfs/shapefile.pdf`

5. Now right-click on `PubSchools.geojson` and again choose Layer Properties. You will see that this is a GeoJSON dataset and that it has 287 point features. GeoJSON (`http://geojson.org/`) is an open standard format that stores spatial data as JavaScript. It is commonly used for displaying vector data in web maps.

6. Expand the `transportation.gpkg` database. Notice that is has a line geometry icon .

7. Using what you know, open the Layers Properties for the SDOT_StateRoutes layer. The Storage reads GPKG. This is a fairly new file format known as a GeoPackage. It is a spatial database format based on the FOSS SQLite database. It can be used on any operating system. You can see this layer does indeed have line geometry with 122 features. Close the Layer Properties window.

Paul Ramsey authored a very helpful blog entitled, *Beyond the Shapefile with File Geodatabase and GeoPackage*, comparing the limitations of the Shapefile with Esri File Geodatabases and GeoPackages. You can find the article here: `https://carto.com/blog/inside/fgdb-gpkg/`

8. Again right-click on `Hawaii_Counties.shp` but this time choose Add Layer to Project. The counties layer will be added to QGIS. It will appear on the map canvas and as a layer in the Layers Panel (Figure 1.6, on the next page).

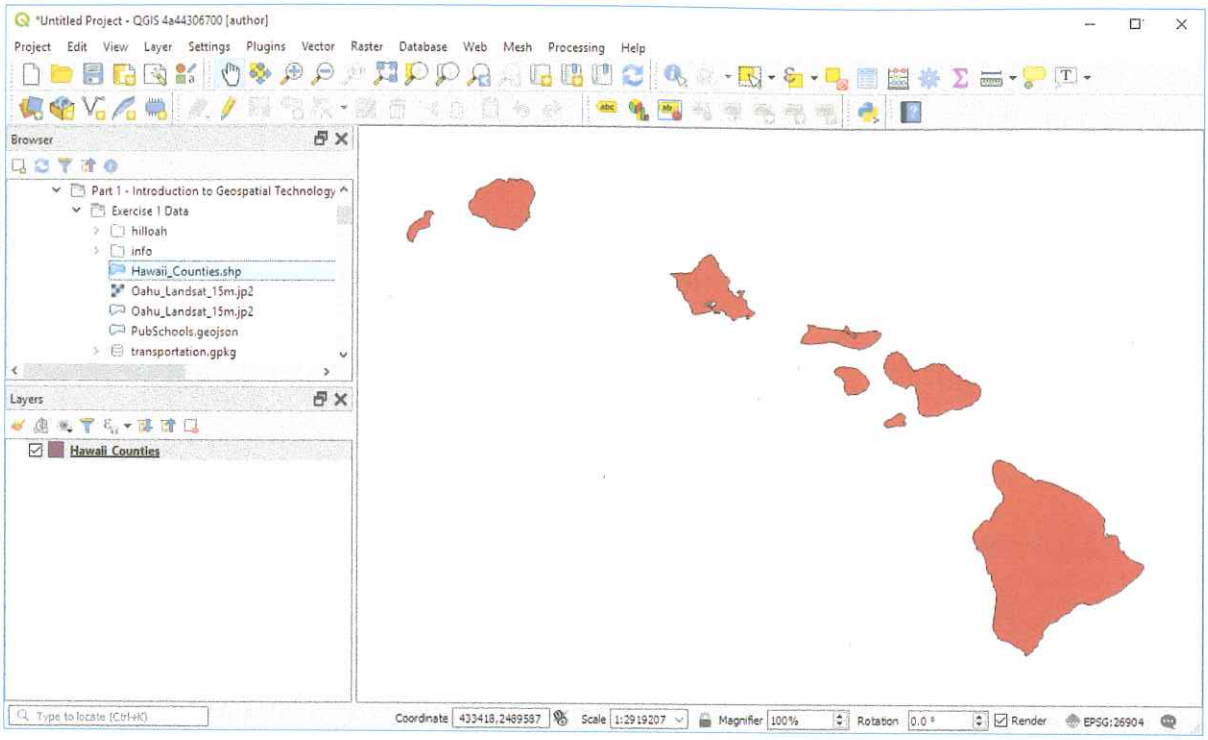
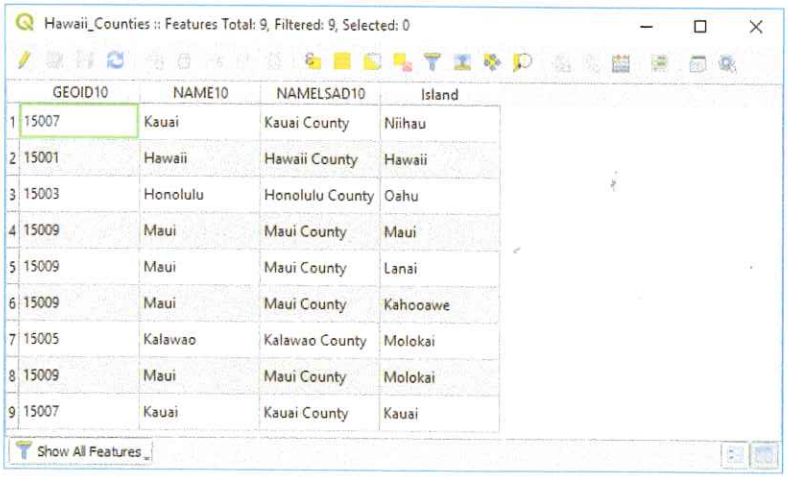


Figure 1.6: First Layer in QGIS

9. Now you will switch your attention to the Layers Panel. Right-click on the Hawaii_counties layer and choose Open Attribute Table from the context menu. The attribute table can also be opened by clicking the Open Attribute Table button or by clicking F6. The table opens showing you the other component of the data model, the attributes (Figure 1.7). Each row corresponds to one polygon. If you recall from exploring this dataset with the Browser Panel, it has 9 features (9 polygons). The attribute table has 9 corresponding records. The columns are things we know about the polygons. There are columns with the County name (NAMELSAD10) and with the Island name Island). Close the table when done.



	GEOID10	NAME10	NAMELSAD10	Island
1	15007	Kauai	Kauai County	Niihau
2	15001	Hawaii	Hawaii County	Hawaii
3	15003	Honolulu	Honolulu County	Oahu
4	15009	Maui	Maui County	Maui
5	15009	Maui	Maui County	Lanai
6	15009	Maui	Maui County	Kahoolawe
7	15005	Kalawao	Kalawao County	Molokai
8	15009	Maui	Maui County	Molokai
9	15007	Kauai	Kauai County	Kauai

Figure 1.7: Attribute Table

10. Using what you know add the Oahu_Landsat_15m.jp2 dataset (with the raster icon) to the map canvas.

11. Right-click on the layer in the Layers Panel and choose Zoom to Layer from the context menu. This is an example of a raster dataset. Like a photograph, it is composed of cells or pixels. This raster is a satellite image of the island of Oahu, Hawaii (Figure 1.8). Close QGIS when done.

Figure 1.8: Oahu Satellite Image

1.5 Task 3 - Working with the Data Source Manager

Now that you know how geospatial datasets are stored on your computer, you will learn more about the data they contain. This next section will introduce you to working with the QGIS Data Source Manager.

1. Click Start | All Programs | QGIS | QGIS 3.x.

The QGIS interface is a little cluttered by default, so let's arrange the panels so the Layers panel and the Browser panel take up the same space.

Panels can be docked and undocked from the QGIS window. To undock a panel, click and drag the panel's top title bar (outlined in Figure 1.9) and drag it away from the sides. When you release your mouse button, the panel will be floating freely.

Figure 1.9: Area to drag when undocking a panel

To dock a floating panel, click and drag the title bar, and drag the panel to the left or right side of QGIS until a rectangle appears underneath the panel. To stack the Layers panel with the Browser panel drag the Layers Panel over the Browser Panel until it is highlighted and drop it. At the bottom of the panel you will now see a tab for each panel. Each are now available with a click of a button, and each panel has more space when you're working with it Docking action is shown in Figure 1.10, on the next page.

With the QGIS interface customized, let's add some data. In Task 2, you learned how to add data via the Browser panel. Here you will learn about the Data Source Manager.

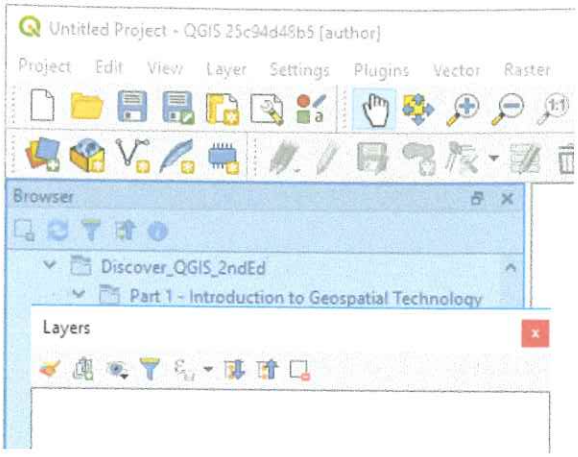


Figure 1.10: Docking the Layers Panel

On the Data Source Manager toolbar click the Open Data Source Manager button . The Data Source Manager window opens. Along the left side are tabs for the Vector and Raster data models along with database and web server connections. The top most tab is the Browser. With that selected the right side of the window shows a Browser window similar to the Browser Panel you have been working with so far (see Figure 1.11, on the facing page).

2. Click the Vector tab Vector

3. The right portion of the window will switch to a dialog used for adding vector data to QGIS. Next you will add one of the ESRI shapefiles. Since this is a file based dataset you will keep the Source type File which is the default. Then click the Browse button. (Figure 1.12, on the next page).

4. The Open an OGR Supported Vector Layer window opens. The window defaults to all files. Click the All files dropdown box and change it to ESRI Shapefiles (shown in Figure 1.13, on page 28). Take a moment to peruse the other formats supported.

NOTE: OGR is a FOSS4G project with the sole purpose of reading and writing geospatial vector data files. You can see which formats it supports here: `https://www.gdal.org/ogr_formats.html`

5. Once you are finished exploring, make sure it is still set to ESRI Shapefiles. This filters what you see in the exercise folder so that you only see the shapefile(s).

6. Select `Hawaii_Counties.shp` and click Open (see Figure 1.14, on page 28).

7. Now back at the Data Source Manager window, click Add to add the data to QGIS and Close to dismiss the Data Source Manager window.

8. You will now see Hawaii_Counties in the Layers Panel and the map features displayed in the map canvas. Vector GIS layers will come in with random colors. You will learn how to change layer styling in a future exercise.

9. In Task 1 you opened the attribute table. Another way to interact with both the spatial features and the attributes is the Identify Features button.

10. Click the Identify Features button

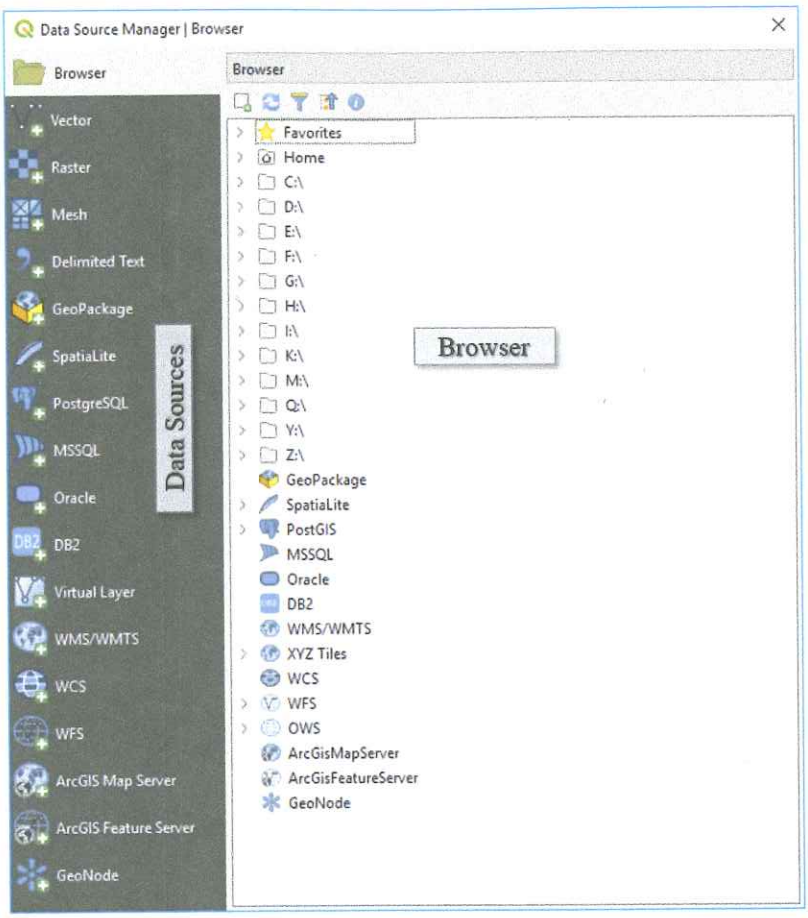


Figure 1.11: Data Source Manager

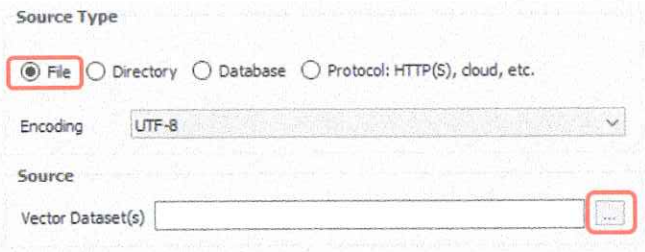


Figure 1.12: Add Vector Layer

11. Click on one of the features on the map. The Identify Results panel (Figure 1.15, on the following page) shows you the attributes for the feature you clicked on. *Note:* The Identify Results panel may initially be docked or floating. You can position it as needed.

Now you will learn how to add Raster data with the Data Source Manager .

12. Open the Data Source Manager .

13. Click the Raster tab Raster .

14. Keep the source type as File and browse to your `Exercise_1_Data` folder.

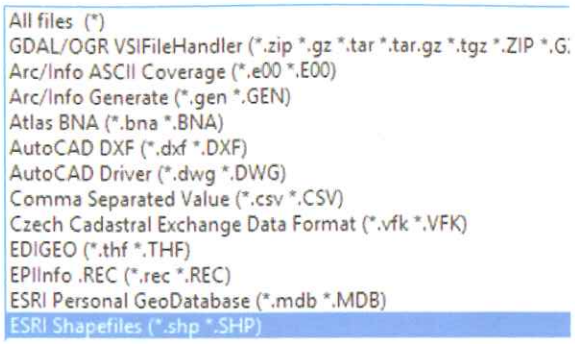


Figure 1.13: OGR Supported Vector Formats

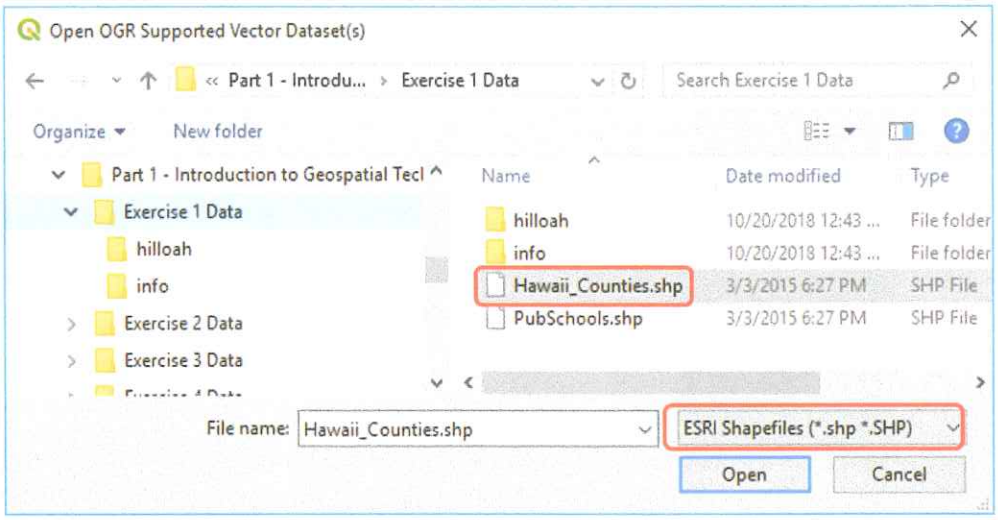


Figure 1.14: Open an OGR Supported Vector Layer

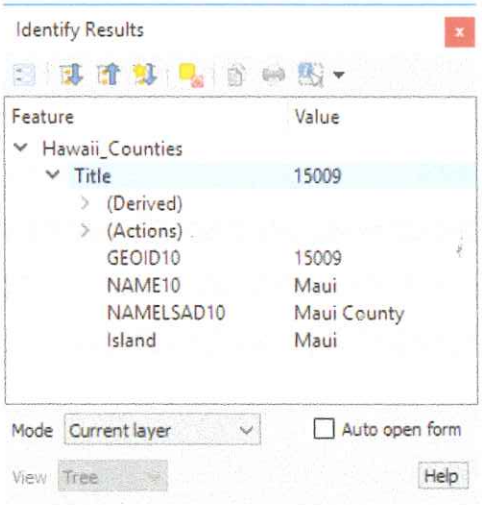


Figure 1.15: Identify Results

15. The Open a GDAL Supported Raster Data Source window opens. This is a very similar workflow to adding

vector data.

Whereas QGIS uses OGR to open vector data files, here it uses another FOSS4G software library called GDAL. GDAL is used for reading and writing raster datasets. You can review the supported GDAL raster formats here: `https://www.gdal.org/formats_list.html`

16. The windows raster data filter is set to All Files by default, so you see the entire contents of the folder.

17. Set the filter to ERDAS JPEG2000. Also, note how many formats it will read! There are many more raster file types than vector. Once you have set the filter you'll see the one dataset: `Oahu_Landsat_15m.jp2` (Figure 1.16).

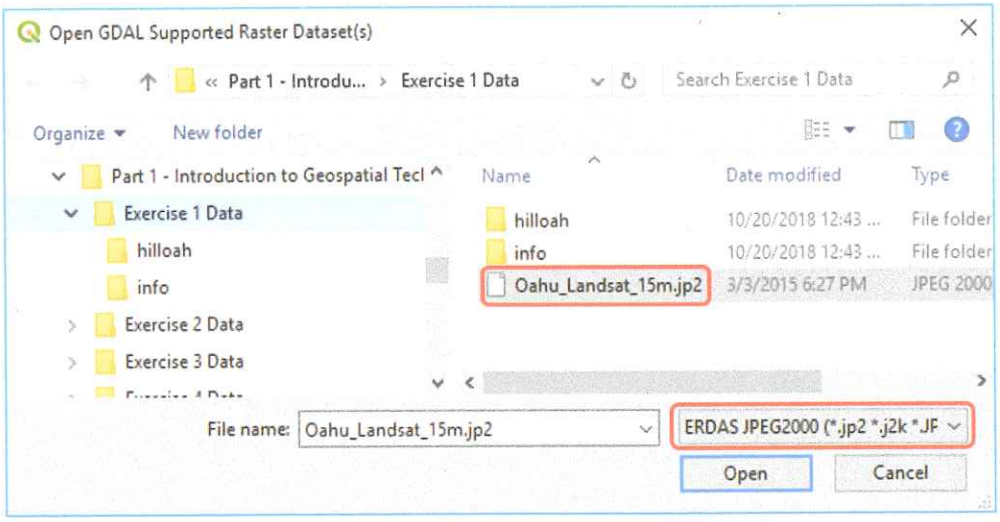


Figure 1.16: Open a GDAL Supported Raster Data Source

18. Select the `Oahu_Landsat_15m.jp2` raster dataset and click Open and then back at the Data Source Manager click Add and Close.

19. This dataset only covers a portion of Hawaii—just the island of Oahu. Right-click on the Oahu_Landsat_15m dataset in the Layers Panel and choose Zoom to Layer to zoom to the spatial extent of this raster (Figure 1.17, on the following page).

You may notice two folders in the exercise data folder that we have not discussed yet. One is named `hilloah` and the other `info`. Together, these combine to make another geospatial raster dataset format named a GRID. The info folder holds the attributes and always has the name "info". The other folder is the layer name and contains the spatial data. Let's add a GRID raster to our map.

20. Open the Data Source Manager and click the Raster tab again.

21. Set the filter to Arc/Info Binary Grid. Double click the `hilloah` folder to enter it. Select the `hdr.adf` file and click Open. Back at the Data Source Manager click Add and Close to add the raster to QGIS (Figure 1.18, on the next page).

22. This raster is a hillshade image of Oahu and it represents the terrain.

Data is often stored deep inside a series of folders. It is often tedious and time consuming to navigate deep inside the folders to gain access to the data. Favorites provide a way to create a shortcut directly to any folder so that you have one-click access. Let's create a favorite to our exercise folder for practice.

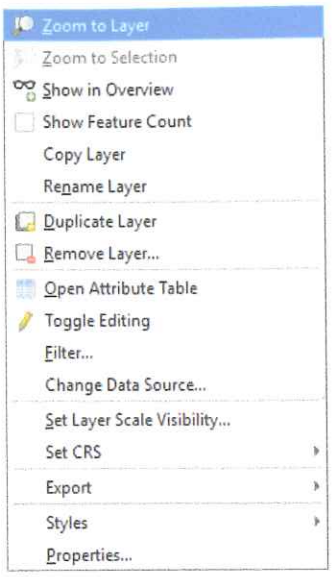


Figure 1.17: Zoom to Layer Extent

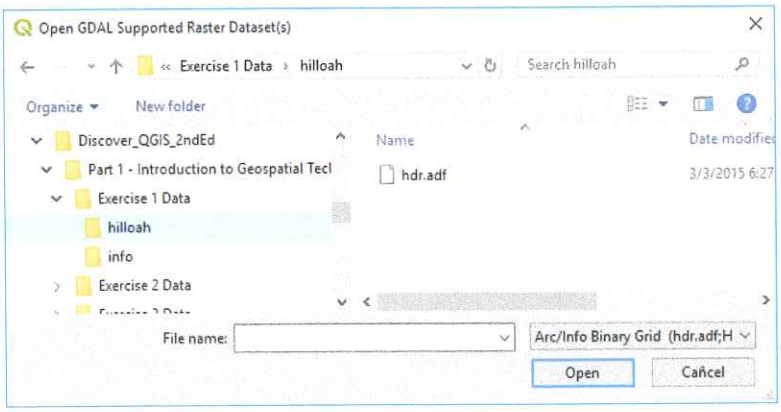


Figure 1.18: Adding a GRID

23. Look at the Browser Panel and scroll to the top. Note that there is a Favorites item. You can identify folders or locations as being favorites in order for them to appear here.

24. Navigate to the `Exercise_1_Data` folder in the Browser Panel. Right-click on it and choose Add as a Favorite (Figure 1.19, on the facing page).

25. Now expand Favorites and you will see your exercise folder listed there. You can remove a favorite anytime by right-clicking on it and choosing Remove favorite.

26. Expand the exercise folder under Favorites to expose the contents. Select `SDOT_StateRoutes.geojson` and drag it onto the map. This is a quick way to add data to your map.

1.6 Conclusion

In this exercise you explored datasets that use the two common geospatial data models: vector and raster. You have also used the Browser panel and the Data Source Manager to add data to QGIS. In future exercises, you will learn how to use QGIS to make maps and perform analyses.

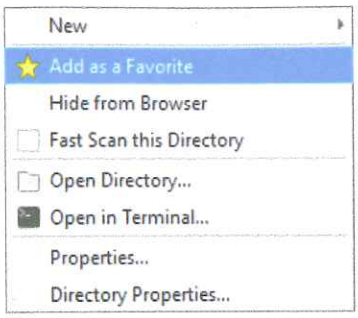


Figure 1.19: Add as a Favorite

1.7 Discussion Questions

1. How can Browser favorites make your workflow more efficient?
2. What are the two main parts of a GIS data model?
3. Name two ways of seeing feature attributes for a vector GIS layer.