Extracted from:

Desktop GIS
Mapping the Planet with Open Source

This PDF file contains pages extracted from Desktop GIS, published by the Pragmatic Bookshelf. For more information or to purchase a paperback or PDF copy, please visit http://www.pragprog.com.

Note: This extract contains some colored text (particularly in code listing). This is available only in online versions of the books. The printed versions are black and white. Pagination might vary between the online and printer versions; the content is otherwise identical.

<div align="right">Chapter 3</div>

Working with Vector Data

In this chapter, we'll start working with vector data (points, lines, polygons) by viewing, editing, and analyzing various datasets. Not only will we view data, but we'll look at tweaking the way data is displayed to make it convey more information at a glance.

3.1 Viewing Data

Viewing data is like the "Hello, World!" application that everyone writes when learning a new programming language. It's the first thing you're going to want to do with any GIS application. Let's start out by seeing what kind of things we can do with vector data using open source GIS software. If you recall Harrison's original project, he first just wanted to view bird locations. We'll take a similar approach and start by viewing some sample vector data.

Viewing data is really more like visualizing the relationships between the features. You can get a lot of information by simply viewing features and applying some special rendering techniques.

When it comes to software, we have a lot of choices for viewing GIS data (see Appendix A, on page 283). As we begin to explore our data, we'll use several different applications to give you a feel for what's available.

Before we can begin, we obviously need some data to work with. If you don't have a shapefile or two handy, you can download[1] a sample dataset and use it to follow along. We will be using this dataset throughout the following chapters when we need to illustrate some basic func-

---

1. http://desktopgisbook.com/sample_data

Joe Asks. . .

What Is a Shapefile?

A *shapefile* stores vector features and their attributes. A given shapefile can contain only one type of feature: points, lines, or polygons.

The term is actually a bit misleading, since a shapefile always consists of at least three separate files. For example, a shapefile named alaska would consist of the following:

- alaska.shp containing the spatial features
- alaska.dbf containing the attributes
- alaska.shx, which is an index file that allows random access to features in the alaska.shp file

In addition to the three main files described here, you might also find alaska.sbx, alaska.sbn, and alaska.qix files. These are additional index files used by some applications. One last file you'll often find associated with a shapefile is a .prj file. This file contains the projection information for the shapefile, including the geodetic datum (for more on datums, see the *Joe Asks. . .* on page 150).

If you are sharing a shapefile with someone, make sure you include at least the .shp, .dbf, and .shx files; otherwise, it will be unusable.

tions or concepts. The dataset includes world borders, cities, and a nice raster image of the earth (which we'll use in a later chapter when we work with rasters).

Choosing a Viewer

Most of the applications in Appendix A, on page 283, that work with vector data go beyond a viewer. Let's use several of them to look at the sample data. Of course, you don't need to use all of them, but following along will help you decide which is best for you. For help on installing any of the applications, see Appendix B, on page 305.

The truth is that nearly all the OSGIS viewers use a similar user interface. If you can use one, you can figure out the others. Let's start by

viewing the world borders data using the User Friendly Desktop Internet GIS, uDig.

Simple Viewing

If you need help installing uDig, take a look at Section B.4, *uDig*, on page 308. OK, let's fire up uDig so we can get a look at that sample data:

- *Linux*: Change to the udig subdirectory, and run udig.

- *Mac OS X*: Double-click the uDig icon in your Applications folder.

- *Windows*: Click the Start button, find the uDig program folder in Program Files, and choose uDig.

When you first start uDig, you are presented with a start-up screen. You can explore the options, but if you are anxious to get busy, click the curved arrow in the upper right of the workspace. This gets us to the business end of uDig.

The uDig workspace isn't much to look at the first time you run it. You'll notice that when uDig starts up, it displays a fairly typical Tip of the Day dialog box. Feel free to click through the tips and see what pearls of wisdom you can find. You can turn off this feature if it bothers you (or you've read them all).

Now let's load the world borders layer to get a feel for how uDig manages layers, as well as the options for symbolizing features. To view the data from our sample dataset, start by clicking the Layer menu and then choosing Add. This opens the Data Sources dialog box, as shown in Figure 3.1, on the next page.

As you can see from the Data Sources dialog box, uDig supports a good selection of formats. Let's start by adding our shapefile of all the countries in the world. Since this is a file-based data store,[2] we choose Files from the Data Sources dialog box and click Next, which opens a file selection dialog box from which we can choose our shapefile. We navigate to the directory containing the shapefile (in this case world_borders.shp) and click Open (or whatever the standard dialog box calls it on your platform). This loads the shapefile into uDig and displays it, as shown in Figure 3.2, on page 44. We've closed the Web

---

2. This is a fancy way to say a GIS data file on your disk drive as opposed to web-accessible or spatial database data.
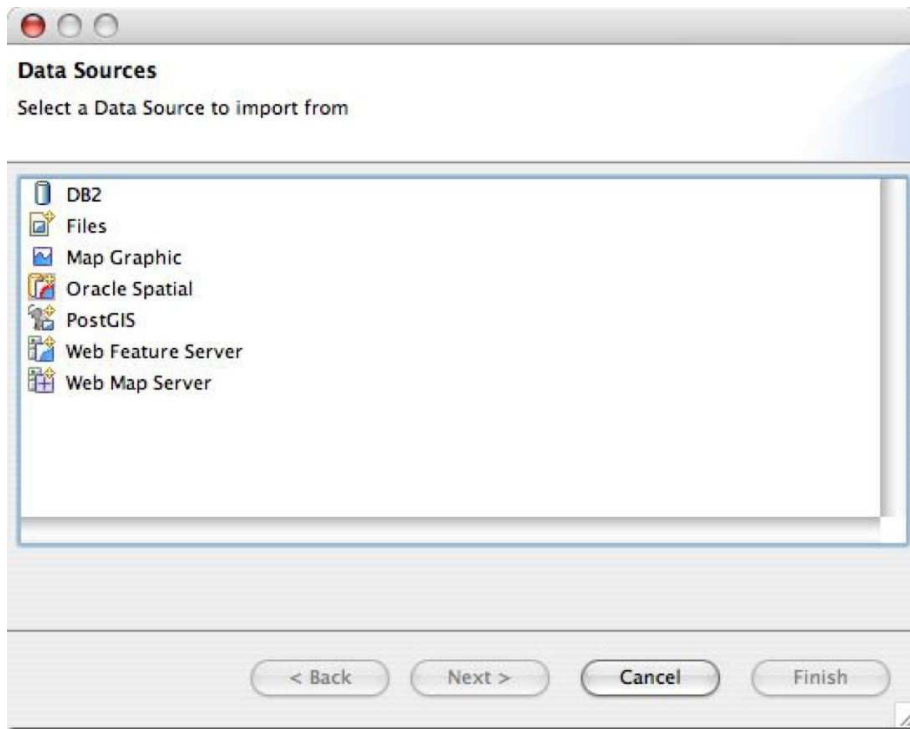
Figure 3.1: uDig Data Sources Dialog

Browser tab to maximize the map area and still give you a feel for the entire interface.

If you are feeling adventurous, go ahead and load the cities layer as well, using the same process.

Moving Around

If you're following along, you should be looking at the countries of the world. Take a look at New Zealand—it's pretty small. This is where navigation tools come into play. Every GIS application, whether it be on the Web or the desktop, has a way to navigate around the map. uDig, of course, supports the usual zoom/pan/identify functions common to all applications.

Let's get a closer look at New Zealand. Select the Adjust Current Zoom tool from the toolbar. It's the magnifying glass with the drop-down caret next to it. If you are unsure which tool it is, hover the mouse for a few
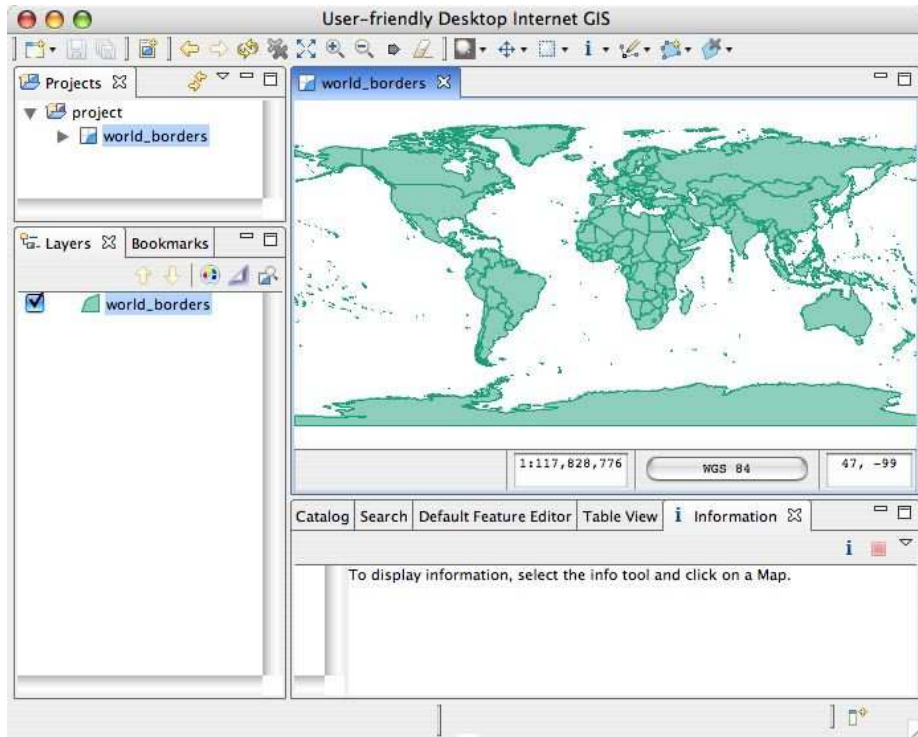
Figure 3.2: uDig displaying world borders

seconds, and you'll get a tooltip to help you out. Find New Zealand, drag a box around it, and then release the mouse. uDig will zoom the view to cover the region of the box. You now have a better view of New Zealand. You can continue to zoom in as much as you like by dragging boxes with the mouse.

So now that we've zoomed into the gnat's eyebrow, we need to determine how to get back out. There are a couple of ways do it. First we can go back to the full view (extent) by using the Zoom to Layers tool in the toolbar above the layer list. This will zoom to include all layers on the map. This may not be what we want if we just want to look at the full view of the world_borders layer. To zoom to just its extent, right-click world_borders in the layer list, and choose Zoom to Layer.

We can also zoom out incrementally by using the Zoom Out tool on the main toolbar. Unlike the Adjust Current Zoom tool, this tool is a one-shot

affair—you don't interact with the map when using it. With each click, the map is zoomed out by a fixed amount. By now you've probably noticed its cohort, the Zoom In tool. Clicking it zooms the map in by a fixed amount.

One last way we can navigate the map is by panning. To pan the map, select the Pan Map View tool (actually both this and the Adjust Current Zoom buttons are tool groups but contain only one tool) from the main toolbar, and drag with the mouse to change the view. You can pan all around the map using this method.

By combining the pan and zoom tools, you can pretty much navigate around the world until your heart's content. You can change the map view in other ways, but we'll leave that for you to discover.

3.2   Rendering a Story

Now it's time to change the way the world looks. This is known as *symbolizing* your data, and you can do it in several ways. Are you happy with the colors uDig chose for the layers? If you are like most people, you have preferences when it comes to these things, and my guess is you're going to want to change the way things look. The simplest, of course, is just a single color for all features, and this is in fact the way all vector layers look when first added to uDig. In uDig you can change the outline color, fill, and marker symbols using the Style Editor. The Style Editor also allows you to turn on labeling and set the maximum and minimum scales at which the layer is displayed.

Our friend Harrison, being the inquisitive sort, quickly decides he wants to be able to tell at a glance where the most populated countries are in the world. Well, we are in luck. Our world_borders layer just happens to have an attribute named POP_CNTRY that contains the population of each country. To make Harrison happy, we can use what's termed a *class break method* to symbolize the data. The Style Editor has a Theme panel to classify the layer based on an attribute. In Figure 3.3, on the next page, we can see the settings we can use to classify the world boundaries based on population. We set up ten equal interval classes based on the range of population values. The less populated countries will be colored in light blues or greens, while the countries with the largest populations will be rendered in a shade of purple. We could also do a quantile classification, putting roughly the same number of values in each class. But for our purpose, the equal intervals work just fine.

Is It Small or Large Scale?

This can be a constant source of confusion when talking to people about maps, whether they be paper or digital. Let's sort it out now.

A small-scale map covers a large area, whereas a large-scale map covers a smaller area on the ground. The terms *large* and *small* are based on the representative fraction that shows the relationship of one unit on the map to one unit on the ground. A map scale of 1:8,000,000 is smaller than one of 1:24,000 since 1/8,000,000 is a smaller fraction than 1/24,000.

Simple enough. If you ever get confused, just think in terms of fractions, and you'll be able to sort out the small from the large.
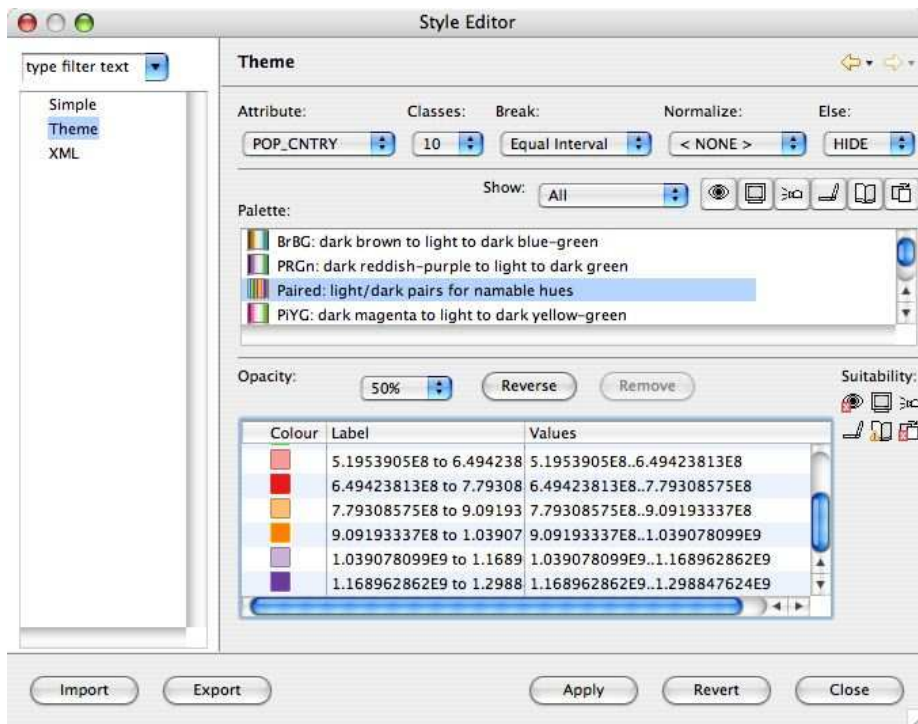

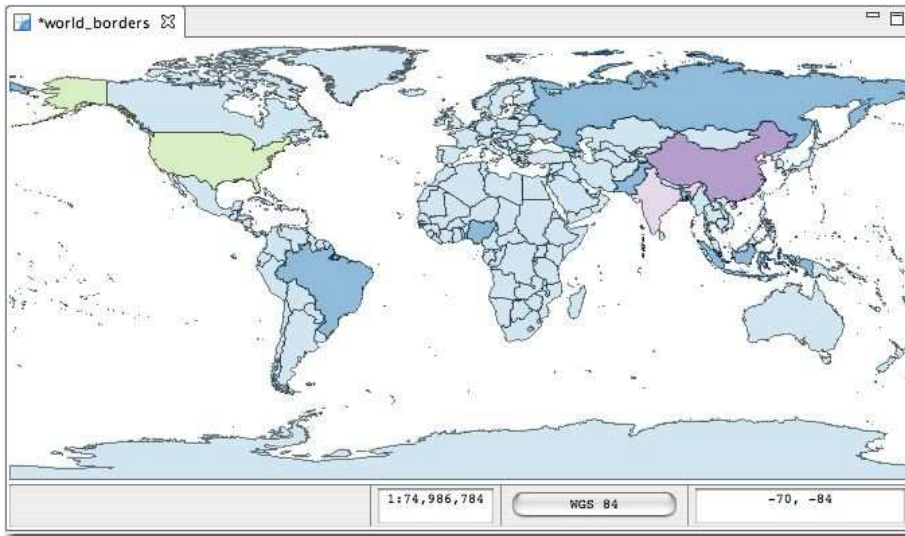Figure 3.3: Classifying countries by population

Figure 3.4: Countries classified by population

The results of the classification are shown in Figure 3.4. As you might have suspected, China and India are rendered as the most populous, followed by the United States. We could refine our classification to get a finer-grained view of population by changing the number of classes or the method. This is a common way to render data to make it tell a story. Some of the other OSGIS applications we will look later offer even more ways to symbolize your data.

3.3 Looking at Attribute Data

In the previous section, you might have wondered how we knew about the POP_CNTRY field in the world_borders shapefile. Well, there are a number of ways to examine the attribute data associated with a layer. One of the quickest ways is with ogrinfo, a utility that is part of the GDAL/OGR suite. We'll take a more detailed look at ogrinfo and friends later in Section 11.2, *Using GDAL and OGR*, on page 195. If ogrinfo is not already on your system, see Section B.7, *FWTools*, on page 310 for information on installing FWTools.[3] Then you can open a command shell on your

---

3.   FWTools is a suite of tools that contains many applications, including ogrinfo.

system and do the following:

```
$ ogrinfo -so -al world_borders.shp
INFO: Open of `world_borders.shp'
      using driver `ESRI Shapefile' successful.

Layer name: world_borders
Geometry: Polygon
Feature Count: 3784
Extent: (-180.000000, -90.000000) - (180.000000, 83.623596)
Layer SRS WKT:
GEOGCS["WGS 84",
    DATUM["WGS_1984",
        SPHEROID["WGS 84",6378137,298.257223563,
            AUTHORITY["EPSG","7030"]],
        AUTHORITY["EPSG","6326"]],
    PRIMEM["Greenwich",0,
        AUTHORITY["EPSG","8901"]],
    UNIT["degree",0.01745329251994328,
        AUTHORITY["EPSG","9122"]],
    AUTHORITY["EPSG","4326"]]
CAT: Real (16.0)
FIPS_CNTRY: String (80.0)
CNTRY_NAME: String (80.0)
AREA: Real (15.2)
POP_CNTRY: Real (15.2)
```

Near the end of the output you'll find the fields included in the dataset. Note that POP_CNTRY is listed last, and the output indicates that it is a numeric field.

Of course, all the desktop GIS applications provide a way to not only determine which fields are in a dataset but actually view the data itself. In uDig we just click the Table View tab below the map view, and we get a nicely formatted view of the data, as shown in Figure 3.5, on the next page. We'll examine working with data in other applications in a bit.

Viewing the attribute table is good for just browsing around. Let's look at some more advanced ways to view and render our data.

3.4 Advanced Viewing and Rendering

Harrison has some more bird-sighting data he collected in his travels. He wants to view the sightings in a number of ways, including by species and number of birds per site. This will allow him to quickly identify where he saw individual species and large groups of the same species. Fortunately, there are some advanced rendering techniques that can help him out.

The Pragmatic Bookshelf

The Pragmatic Bookshelf features books written by developers for developers. The titles continue the well-known Pragmatic Programmer style, and continue to garner awards and rave reviews. As development gets more and more difficult, the Pragmatic Programmers will be there with more titles and products to help you stay on top of your game.

Visit Us Online

Desktop GIS's Home Page
http://pragprog.com/titles/gsdgis
Source code from this book, errata, and other resources. Come give us feedback, too!

Register for Updates
http://pragprog.com/updates
Be notified when updates and new books become available.

Join the Community
http://pragprog.com/community
Read our weblogs, join our online discussions, participate in our mailing list, interact with our wiki, and benefit from the experience of other Pragmatic Programmers.

New and Noteworthy
http://pragprog.com/news
Check out the latest pragmatic developments in the news.

Buy the Book

If you liked this PDF, perhaps you'd like to have a paper copy of the book. It's available for purchase at our store: pragprog.com/titles/gsdgis.

Contact Us

Phone Orders:	1-800-699-PROG (+1 919 847 3884)
Online Orders:	www.pragprog.com/catalog
Customer Service:	orders@pragprog.com
Non-English Versions:	translations@pragprog.com
Pragmatic Teaching:	academic@pragprog.com
Author Proposals:	proposals@pragprog.com