GST 102: Spatial Analysis Lab Series

Lab 2: Introduction to Geospatial Analysis

Document Version: 2013-07-31 (Beta)

Organization: Del Mar College
Author: Richard Smith

The development of this document is funded by the Department of Labor (DOL) Trade Adjustment Assistance Community College and Career Training (TAACCCT) Grant No. TC-22525-11-60-A-48; The National Information Security, Geospatial Technologies Consortium (NISGTC) is an entity of Collin College of Texas, Bellevue College of Washington, Bunker Hill Community College of Massachusetts, Del Mar College of Texas, Moraine Valley Community College of Illinois, Rio Salado College of Arizona, and Salt Lake Community College of Utah.

Contents

Introduction

This lab is part of a series of lab exercises designed through a grant initiative by the National Information, Security & Geospatial Technologies Consortium (NISGTC), funded by the United States Department of Labor in partnership with the Department of Education under the Trade Adjustment Assistance Community College and Career Training Grant Program (TAACCCT).

GIS data comes in many formats. As you collect data from various sources on the Internet, you will find that not all data is perfect and may require some post-processing or normalizing for it to be useful and meaningful. In most cases, the data you acquire will not be spatially enabled, in other words, there is no spatial component to the data. Some datasets may be large and come in a comma-delimited format (.csv) or as database files (.dbf). These files can be converted or changed in ArcCatalog to a format that can be used in ArcGIS.

Your instructor may require that you provide screen captures, exported files and/or a copy of this lab with questions answered. Please check with your instructor for requirements specific to your class.

This lab includes the following tasks:

1. Data Exploration and Joins
2. Creating Scatter Plots, Graphs, and Charts

Objective: Explore and Understand Geospatial Data Models

In this lab, we will look at data and decide if it needs post-processing or any editing before we import it into ArcGIS. The data will be joined to a shapefile.

Scatter Plot– a plot that displays multiple datasets to analyze patterns and relationships

Lab Settings

Required Virtual Machines and Applications

Windows Machine User Account	Train
Windows Machine User Password	Train1ng$

1 GIS Data Structures

Understanding how to make data useful in a GIS, is a vital part of working in a geospatial realm. This lab provides practice using and manipulating data files much like those you might commonly find when collecting data for a spatial analysis.

The *Lab 2* folder contains three files that will allow you to analyze the spatial distribution of annual CO_2 emissions and the percentage of renewable energy use per country.

The data in these files require no processing, but take note of the file extensions. In order to give the data a spatial component we need to join it to a shapefile. In order to join any other file to a shapefile, we must ensure that there are fields in common.

1. Log into the computer, using the information provided in the Lab Settings section.
2. Copy the data for this lab from its Lab 2 folder location on the lab machine into your folder: *C:\GST 102*
3. Click **Start->All Programs->ArcGIS->ArcCatalog 10.1** this will open ArcCatalog.
4. Find and expand the GST *102/Lab 2* folder in the catalog tree. Click on the *Lab 2* folder to display the contents in the display window. When you click on a folder, the display window will display the contents of the folder.
5. Preview the data in ArcCatalog by selecting a file in the catalog tree such as **RenewableEnergy_Percentages.csv** and the file's contents will open in the display window.
6. Select the **Preview** tab next to the contents tab and a table will appear.
    a. To preview excel files(.xls) we have to go into the file and select the data sheet.
    b. Within the **CO2_Readings_World.xls** file there is only one sheet contained by the file and it will appear as **Sheet1$** when the CO2_Readings_World.xls is expanded.
7. When looking at the shapefile **World_Countries.shp** we can preview either the geography or the table of the shapefile as it contains both. To change the preview view, select either Geography or Table from the dropdown box in the bottom left corner of the display window.
8. Exploring the data provided for the lab we can see that there is a common field in the following files:
    a. CO2_Readings_World.xls
    b. World_Countries.shp
    c. RenewableEnergy_Percentages.csv

Question 1: What is the common field in these files?

9. In ArcCatalog right-click on **World_Countries.shp** and under **Properties** select the **XY Coordinate system** tab. We are going to set the coordinate system to **GCS_WGS_1984**. The GCS prefix represents Geographic Coordinate system. Expand the Geographic Coordinate System and scroll down to **World**. Expand World and scroll down to select **WGS 1984**. Click **Apply** and then **OK**.
10. Next, we will export CO2_Readings_World.xls to the dBase Single format:
    a. In ArcCatalog, expand **CO2_Readings_World.xls** using the plus next to the file. There will be a Sheet1$ within the file.
    b. Right-click on that file and select **Export -> to dBASE (single).** This should open the dialog box shown in Figure 1.
    c. Browse to and select an output location for your file (probably your C:\GST 102\Lab2 folder.
    d. Name your output table `CO2_World` and press **OK** to create your **CO2_World.dbf** file.

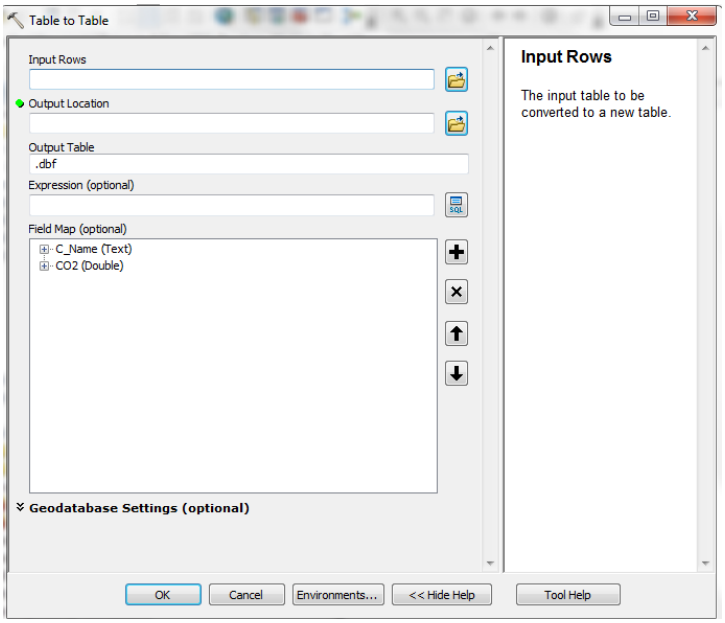

Figure 1. Dialog box for creating a database from a data file.

11. Open a blank map in ArcMap and add CO2_World.dbf file and your newly-projected World_Countries.shp file to the document by dragging and dropping them from the Catalog tab into the Table of Contents.
12. Join the **CO2_World.dbf** to the **World_Countries.shp** by right-clicking on the **World_countries** layer and select **Joins and Relates->Join**.
13. Where it asks, *What do you want to join from this Layer?* **Join Attributes from a table** should be selected on the dropdown box. In the next dropdown, choose **C_name** as the field in the *Choose the field in this layer that the join will be based on;* this field represents the country name.
14. In the next dropdown, select the **CO2_world.dbf** file you created.
15. In the following dropdown, select **C_name**, this means that you are joining the attributes in **World-Countries.shp** to **CO2_world.dbf** using the common field **C_name**.

16. In the **Join Options**, select **Keep only matching records**. This will keep the records for which there is a matching value; all other records will be hidden.
17. We will now symbolize the map using the **Equal Interval, Defined Interval, Quantile, Natural Breaks, Geometrical Interval, and Standard Deviation** by right-clicking on the **World Countries** layer and selecting **Properties**.
18. Select the **Symbology** tab. On the left there is a box that has the heading ***Show:*** select the **Quantities** option. In the **Fields** box select **CO2** for the Value dropdown. This is the field you have joined to the countries map. Do not select any normalization.
19. In the classification box, click the **Classify** button. There are several methods of breaking up the data and displaying it on the map. These include **Equal Interval, Defined Interval, Quantile, Natural Breaks, Geometrical Interval, and Standard Deviation**. Select these methods one at a time and note the ways in which data display changes after clicking OK and Apply.
20. Follow Steps 10-19 to export **RenewableEnergy_Percentages.csv** to a .dbf format and join with **World-Countries.shp.**

2 Scatter Plots, Graphs, and Charts

1. Open a blank map in ArcMap and add the **State_2010Census_DP1** layer to the map by clicking the **Add Data** button and browsing to the file.
2. To create a scatter plot in ArcGIS, click on the **View- > Graphs -> Create Graph**. The Create Graph Wizard should open (see Figure 2 below)

   Use these settings:
   a. graph type: **`Scatter Plot`**
   b. layer/table: State_2010Census_DP1
   c. Y field: DP0010039 (**`Number of females per state`**)
   d. X field: DP0010020 (**`Number of males per state`**)

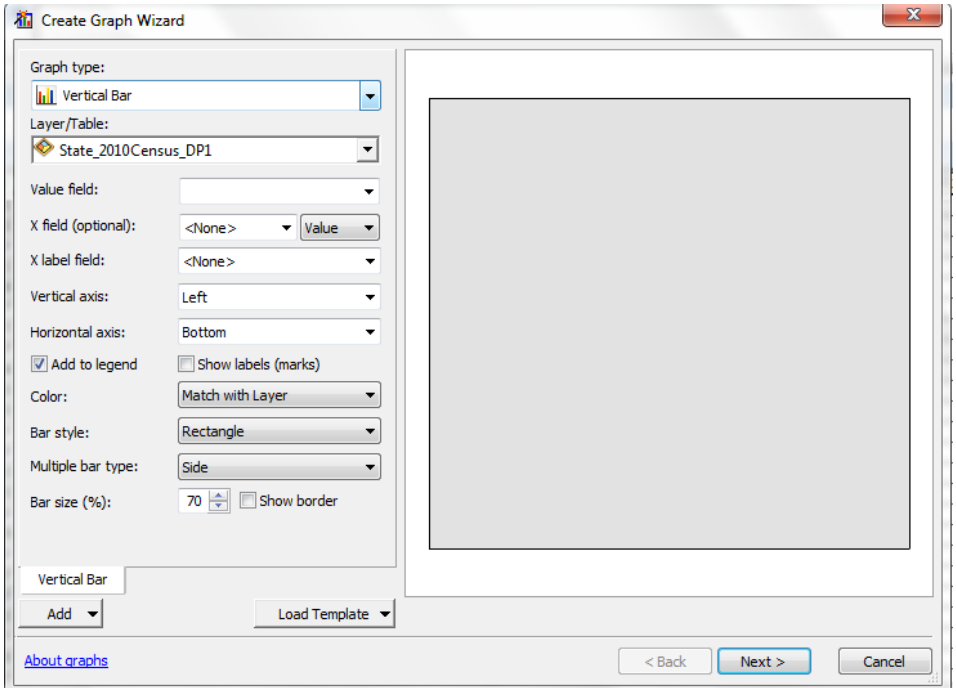

Figure 2. Create Graph Wizard window

3. Once you have assigned the fields to the axis, click **Next**.
4. For the Title, type **`Females Vs. Males per State`**. In the Axis properties, select the Left tab and type **`Females`**, then select the Bottom tab and type **`Males`**.

5. Click **Finish**, and the graph is complete. You should be able to see a linear relationship between the datasets as is shown in Figure 3 below.

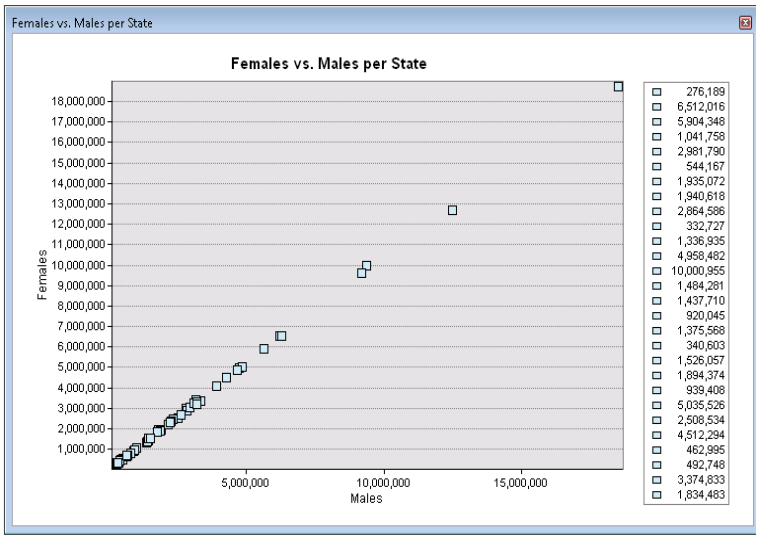

Figure 3. Females vs. Males final graph

6. If you want to view more specific information on the data, you can right-click on a data point on the graph and select **Identify**.

Conclusion

In this lab, you learned to join spatially enabled data. Once that was complete, you were able to analyze data within a scatter plot. These are important techniques within spatial data analysis.

Discussion Questions

1. What is the use of a join in GIS?
2. When joining tables, would you typically keep all records or keep only the matching records? Why?
3. Describe the use of your scatter plot. Describe some examples of data that could be illustrated with a scatter plot.