

What Do You Want to Know?

Beginning Spatial Analysis

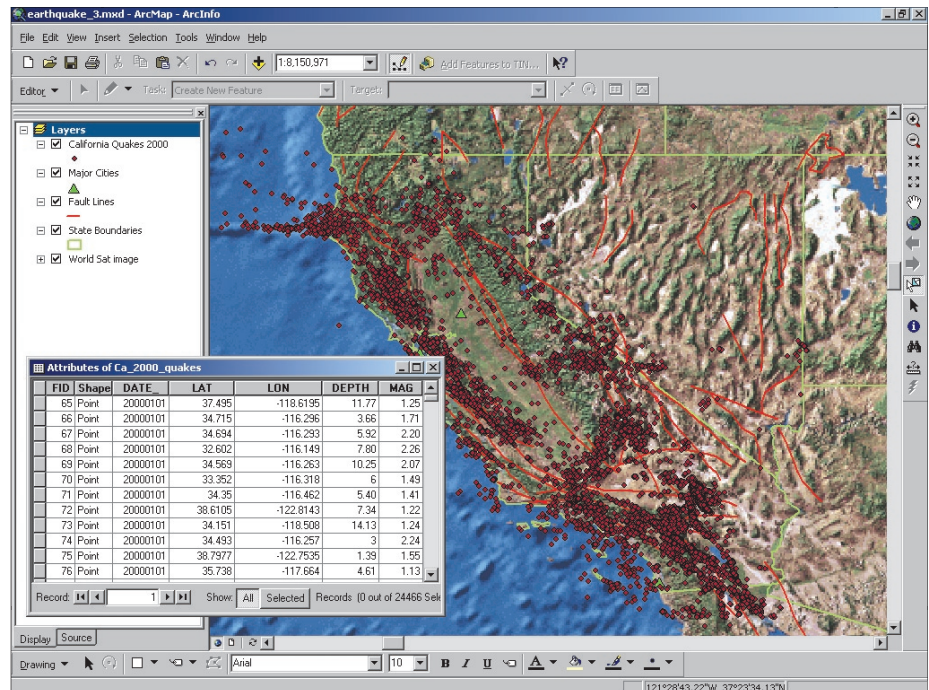
By Monica Pratt, ArcUser Editor

New GIS users who are just getting to know the ArcGIS suite of tools for capturing, managing, and presenting geospatial data sometimes overlook one of the most exciting, powerful, and distinctive capabilities of GIS—spatial analysis. This article introduces new users to some basic spatial analysis concepts and techniques using ArcMap, one of the desktop applications of ArcGIS. Although the ArcGIS Spatial Analyst and ArcGIS Geostatistical Analyst extensions supply additional tools, the basic spatial analysis functionality of ArcGIS is powerful and directly available through tools and menu choices.

Simply put, spatial analysis encompasses all the operations performed on geographic data that add value. GIS is an information system, and like any information system, it is designed to produce answers. Spatial analysis can be used in an inductive, deductive, or normative manner. Inductive analysis studies the data to discern patterns that might be used to support theories or derive general principles. It can be used to help answer questions such as “Why do some areas have especially high rates of a certain type of cancer?” Deductive applications use geospatial data to test existing theories or principles. These applications might be used to test the validity of relationships between risk factors and the occurrence of a type of cancer. Normative uses identify the best solution to spatial-dependent problems such as store siting.

In the guise of analytical cartography, spatial analysis actually predates GIS. However, to the extent that GIS enabled users to rapidly carry out complex operations against large datasets, the real value of spatial analysis was not realized until the advent of GIS. Probably the single most defining characteristic of spatial analysis, as compared with other types of analysis, is its location-dependent nature. New information is generated by identifying patterns that relate data characteristics (i.e., attribute data that is kept in tables) to the data's spatial characteristics (i.e., geographic locations). The results of analysis change when the locations of the subjects under analysis change. GIS also takes into account a related characteristic of geographic data elucidated by geographer and statistician Waldo Tobler. In his First Law of Geography, Tobler stated that things closer together are more similar than things further apart.

Although spatial analysis processes can be very complex, they can also be quite simple and still yield useful results. A key aspect is the use of maps as a visualization tool. The mind



Symbolizing earthquake events using a single point symbol does show where earthquakes have occurred—but not much more. However, the MAG (magnitude) field can be used to filter the data and display it more meaningfully.

is especially adept at visually perceiving both patterns and anomalies. Part of the power of GIS is that it takes advantage of both computer and human information processing. Analyses that employ queries are probably the most easily understood, and this article outlines the process used in this basic type of analysis.

What Do You Want to Know?

The first step in any analysis is to decide what questions you want to answer. The questions you have can be as simple as wanting to know where particular features are located. The answer to one question can often lead to a series of related questions and additional analyses and query results can provide the basis for further analysis.

In some parts of the world, earthquake activity is a topic of abiding interest. California is one of those places. The accompanying illustrations show how ArcMap can be used to explore earthquake data for this area and answer questions such as

“Do higher intensity earthquakes seem to cluster together in time and space?”

“Have these events occurred near major cities?”

“How many people might be affected?”

Where Is the Data?

ArcCatalog, one of the ArcGIS Desktop applications, will help you locate and evaluate data for your analysis. You can use it to connect to the Geography Network or other Internet servers to search for data online in addition to searching folders on your computer and computers on local networks. Once potential datasets are located, ArcCatalog can display geography, table, and metadata views of the data as well as indicate what type of data is present (e.g., shapefiles, rasters). Be sure to evaluate data for its currency, accuracy, and usage restrictions.

The example illustrated required data on seismic activity and data for a basemap showing landform masses and basic political boundaries that would supply a context for the earthquake data. Since part of the analysis included measuring the effects of earthquakes, data providing both the location and population of major cities was also needed.

Government agencies, such as the United States Geological Survey, provide seismic data online at no charge. Data available in ESRI shapefile format can be directly loaded into ArcMap. Data provided in text files that contain a geographic component, such as latitude and

longitude, can be saved as DBF format files, added as tables, and displayed as a layer using Display XY Data functionality. *ESRI Data & Maps*, a multiple CD-ROM dataset that comes with ArcGIS, is a built-in source of basemap data that contains map data for Europe, Canada, Mexico, and the United States and generalized data for the world.

Put It in a Map

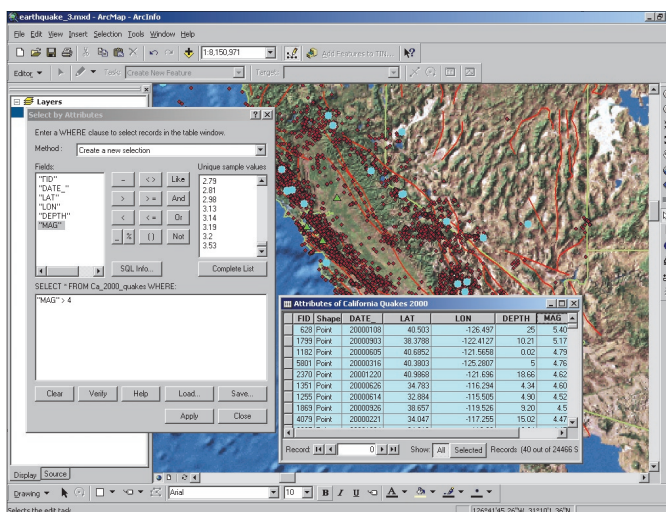
Interrogating a GIS database is more intuitive when data is displayed in a map. ArcMap symbolizes spatial data in layers and organizes those layers in data frames. The ArcMap interface contains tools and menu choices that simplify the process of interactively querying data. Queries can be run against tables or maps. Using the Identify tool, feature attributes can be called up by clicking on map features. Clicking on records in an attribute table will highlight the related features on the map.

Once the necessary data has been evaluated and assembled, the next step is displaying it in a map. The sample earthquake map includes a World WorldSat Color Shaded Relief image used as a background layer and shapefiles showing state outlines, major cities with population data, and fault lines in addition to the earthquake data. Once the data is displayed in the map, we can start to ask some questions about California's earthquakes in 2000.

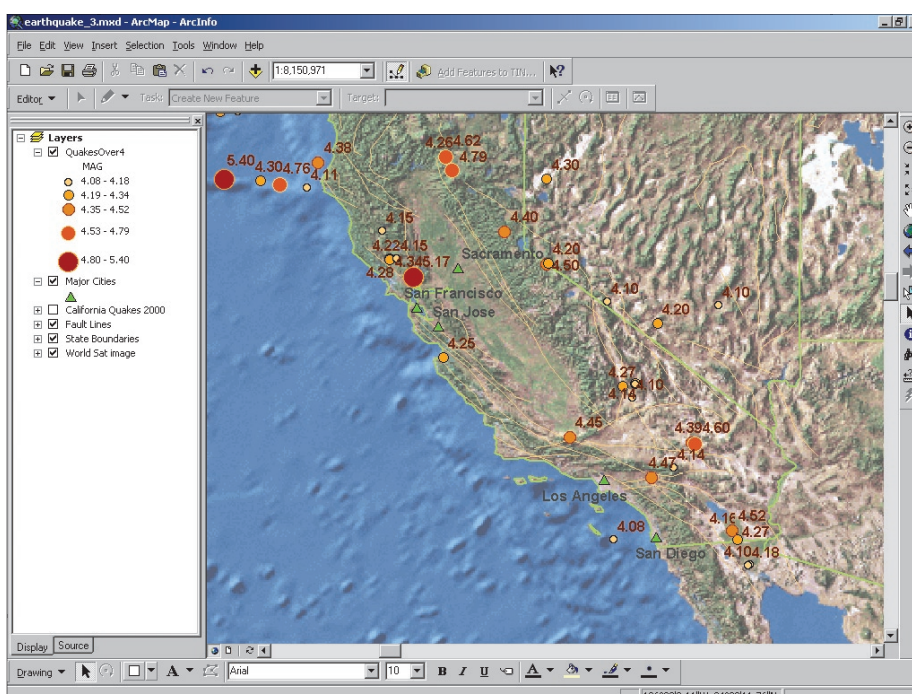
Query the Map

Symbolizing earthquake events using a single point symbol does show where earthquakes have occurred and indicates clusters of events—but not much more. However, when evaluating the seismic data in ArcCatalog, we learned it contains fields giving the magnitude, depth, date, and time of each event in addition to its latitude and longitude. Because an earthquake with a magnitude of 6 is much more unsettling to residents than one that registers 2 on the Richter scale, we might want to include magnitude data in the analysis.

After inspecting the attribute table, it is apparent that many quakes were fairly small-scale events. Querying out quakes with a magnitude of more than 4 will show the more significant events. One way to locate events with magnitudes of more than 4 would be to sort the magnitude field in descending order and hold down the Control key to select all the desired records. For a small table, that might be a good technique but this table has 24,466 records, so using an expression to select features based on attributes is a more efficient method. Choosing Selection, then Select by Attributes, and creating the expression “MAG” > 4 in the query box whittles down the number of events to 40. Exporting the selected features to a new shapefile lets us deal only with those features.



Querying out the larger events using the expression “MAG” > 4 whittles down the number of events to 40 points that can be exported to a separate shapefile.



Symbolizing, labeling, and reconsidering data can lead to new questions and additional analyses.

Modify the Map to Ask New Questions

Now these larger events can be symbolized based on magnitude using both graduated symbol size and color. Labeling each by magnitude and adding labels for the cities lets us easily see which cities were in close proximity to large events. Alternatively, events could be symbolized by time or depth to perhaps uncover other aspects of the data. With the largest quakes easily identified, the fault layer could show if seismic activity is especially high near particular faults. Faults could also be symbolized based on their type. After looking at a single year, you might want to obtain data for other years. Using this iterative process, your reasoning ability, and the tools in ArcGIS, you can better understand the data and the portion of the world it describes.

Resources

ESRI has many resources for learning more about spatial analysis. ESRI Press publishes *The ESRI Guide to GIS Analysis, Volume 1: Geographic Patterns and Relationships* by Andy Mitchell and *Geographic Information—Systems and Science* by Paul A. Longley, Michael F. Goodchild, David J. Maguire, and David W. Rhind. Both are available online from www.esri.com/shop. Paul Longley et al. have also authored an ESRI Virtual Campus course called *Turning Data Into Information*. Another Virtual Campus course that teaches spatial analysis techniques is *Understanding GIS Operations: A Transformational Approach*. Sign up for these and other GIS courses by visiting campus.esri.com.