

Asking the Right Questions

Getting started using queries

By Monica Pratt, ArcUser Editor

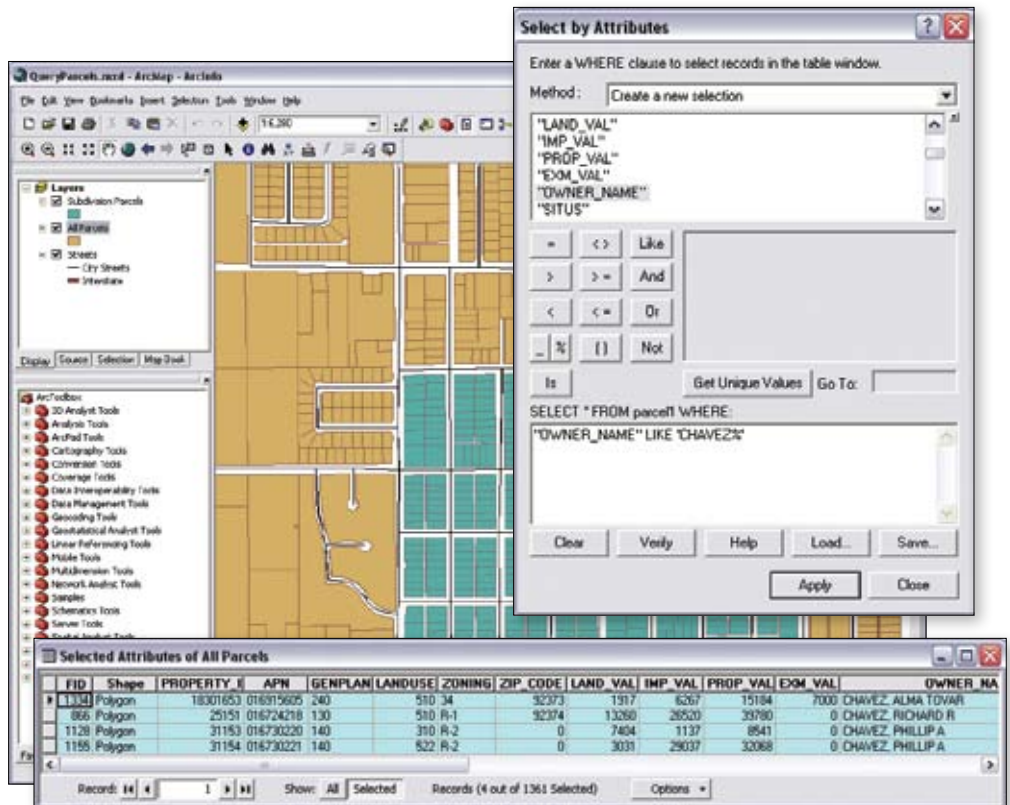
One of the defining characteristics of GIS is the ability to ask questions (and get answers) in a spatial context. These questions or queries are requests for information that are posed in a specific fashion.

Queries are expressed in a formal way using Structured Query Language (SQL). This language is composed of a set of commands and rules that are used to ask questions of a database. SQL, which has been implemented for nearly every DBMS, comes in many subtly different “flavors” depending on the database being queried. However, the details of these differences are not a concern when using the ArcMap application in ArcGIS Desktop. The software detects the characteristics of the underlying database being queried and modifies the interface so SQL statements will be created using the appropriate search terms and information. The user is not required to understand the SQL format for the specific database being queried.

There are two types of queries: attribute and location. Attribute queries ask for information from the tables associated with features or from stand alone tables associated with the GIS. Attributes can be numeric values, text strings, Boolean values (i.e., true or false), or dates. This kind of query is similar to a query made to any database; however, when using a GIS, the answers (i.e., the features related to the records selected by the process) are highlighted on the map as well as in the table.

The answers to spatial queries are derived directly from the location of features on a map. Information about the proximity of one parcel to other parcels or other kinds of features, such as roads, is not contained in an attribute table but is easily learned using a spatial query. For example, you could ask if one or more features are located within a certain distance of other features, are contained by another feature, intersect other features, or possess another of the relationships defined by spatial operators.

GIS is powerful because it can use both attribute and spatial queries to get answers with far less effort or answer questions that would not be practical to answer using any other method.



The *Select by Attributes* dialog box in ArcMap helps build queries by presenting the values and SQL operators available in the underlying database. This query uses *LIKE* to capture all the records with an *OWNER_NAME* that contains the string ‘HAVEZ’.

Building a Basic Query

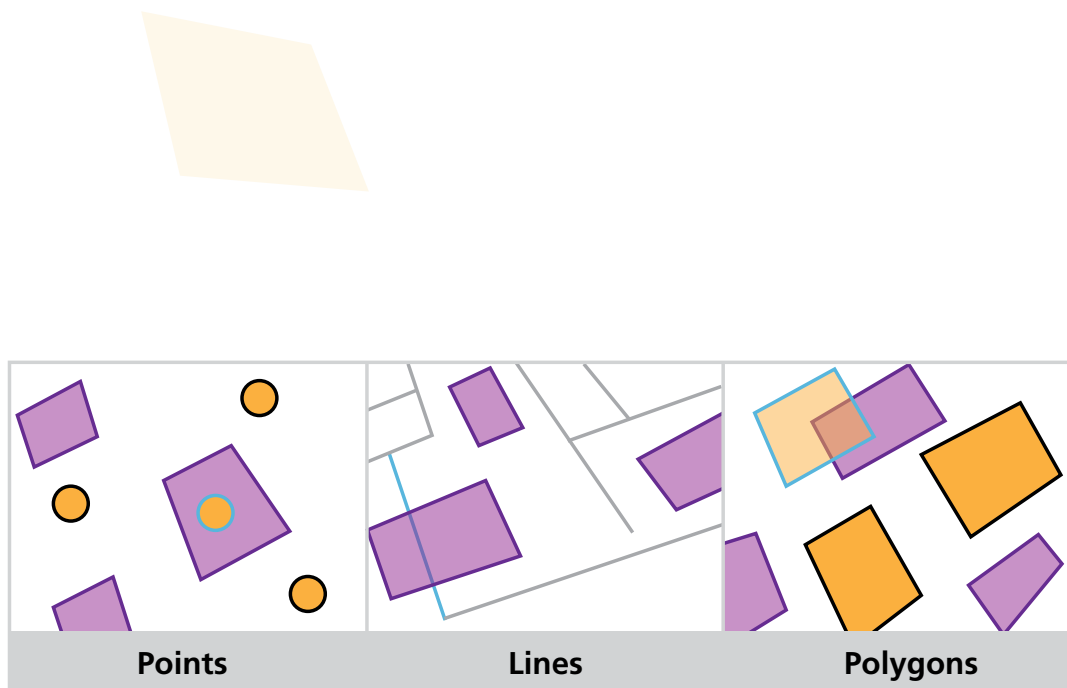
All basic queries have three parts: a source, a filter, and a relationship. This is true of both attribute and spatial queries. The source can be a table or feature class. The filter can be an attribute value or a shape or feature. The relationship between the source and the filter is based on logical, comparison, or spatial operators.

When creating a query, identify which table or feature class will contain the information that answers your question. If the question is, What parcels have a specific commercial zoning? an attribute table for parcel features that contains a field for zoning would likely be a source.

The filter identifies what is different about the desired items, whether those items are table records or features. In the commercial zoning example, that characteristic would be the code C-3, which identifies the commercial zone of interest.

The relationship between the source and filter does the work of finding the record or features. Continuing with the example, the relationship would be that the zoning field contains the value C-3 so the relationship would be “equals” or =.

Relationships are defined using operators. Comparison and logical operators are applied to attribute queries. Comparison operators include =, <>, >, >=, <, and <=. LIKE, AND, OR, and NOT are logical operators. In contrast to the relatively short list of operators for attribute queries, there are more than a dozen types of spatial operators. Intersect, Are Within a Distance Of, Contain, and Are Contained By are some of the most common and useful ones. Spatial queries performed using Select by Location deal with vector data and use a shape as a filter and its relationship with features in the source layer to answer a question.



Spatial queries use the shape and location of one set of features to select another set of features. This illustration shows how a spatial query finds the point, line, or polygon features that intersect with polygon features in another layer.

Using Operators

When selecting by attribute, LIKE and NOT are often used to compare text strings and match patterns, often employing wildcards. Wildcards can replace one character or a group of characters that are unknown. The character used depends on the data source being queried. For personal geodatabases, use a question mark (?) for a single character and an asterisk (*) for a group of characters. For shapefiles, ArcSDE geodatabase feature classes, and other types of data, use an underscore (_) to replace a single unknown character and a percentage sign (%) to replace a group of characters. ArcMap will detect the type of database being queried and adjust the wildcard characters available through the Select by Attribute dialog box.

LIKE is a good operator for finding text strings that contain variant spellings or possibly misspellings, of a text string. Rather than using an equal sign followed by the search term enclosed in quotes, use LIKE and enclose the search term and a wildcard character with quotes.

Like the characters used for wildcards, the syntax used when querying dates depends on the underlying database. ArcMap automatically writes the proper syntax when you double-click on a date value in the Unique Values list of the Select by Attribute dialog box.

The choice of spatial operator (i.e., the relationship tested by the query) depends on the types of features that will be used for the source and filter.

Answering More Complex Questions

To meet multiple search criteria, several attribute queries can be combined using logical operators (such as AND, OR, LIKE, and NOT) to find records based on several criteria in two or more attribute fields.

Remember that OR is the far more flexible and inclusive operator. When using AND, both conditions must be true to return records. For queries that use OR, only one condition must be true to return records.

Ordinarily, queries are evaluated from left to right. However, any portion of the query enclosed in parentheses is evaluated first. The order of operations can be important both in obtaining a valid answer to the question being asked and in optimizing the way a query runs.

A single query such as "LOT_SIZE" >= 1 AND "LOT_SIZE" <= 2 AND "SLOPE" < 5 could be used to locate parcels with lots between one and two acres in size that have a slope of less than 5 percent.

For more complex spatial queries, subqueries can be used to sieve data. The results of one query can be used as the basis for additional queries related to the currently selected features, either selecting from those features, adding to those features, or removing features from the selected set. These operations work in much the same way as AND and OR operators by creating subsets that will be the basis for additional selections.

Working with Both Query Types

Although attribute and spatial queries can work together in locating the desired information, because they are entered in ArcMap in different

dialog boxes (Select by Attribute, Select by Location), these two different kinds of queries must be posed separately.

To expand the previous example (and narrow the search), the desired parcels must have sold within the last year, be zoned C-3, and be located within five miles of a specific parcel in addition to being between one and two acres in size with a slope of less than 5 percent. The date sold and zoning information are also stored in the attribute table but identifying the parcels within five miles of the subject parcel will require a spatial query.

Limiting the parcel search to those that fall into the five-mile buffer around the subject parcel will eliminate many records, so that query is performed first. Using the features selected by location, an attribute query that evaluates the sale date to limit the selection, then the zoning, lot size, and slope will identify features that fulfill all criteria.

Conclusion

Whether the question to answer is simple or complex, break down the question being asked into its query components—source, filter, and relationship—and arrange those components in the order that correctly and efficiently answers that question. Understanding the data distribution in a data source will also help you understand how to best order operations to minimize the number of records that must be searched. For more help on building queries, refer to the ArcGIS Desktop Help Online (available at resources.esri.com/arcgisdesktop) and search for the topic "SQL reference."