



IFLA  
2005  
OSLO

## World Library and Information Congress: 71th IFLA General Conference and Council

"Libraries - A voyage of discovery"

August 14th - 18th 2005, Oslo, Norway

*Conference Programme:*

<http://www.ifla.org/IV/ifla71/Programme.htm>

October 10, 2005

---

**Code Number:**

**201-E**

**Meeting:**

**107 Genealogy and Local History with Geography and Map  
Libraries**

### **Using GIS to Map Genealogical Data: Getting Started**

**Anita K. Oser**

Western Carolina University

USA

---

#### **Abstract:**

Since each of us is born somewhere, lives his/her adult life somewhere, and dies somewhere, genealogical data has a built-in spatial component. It is this spatial component that makes it possible to map the location/residence of an individual during his/her lifetime. The tool that allows you to do this is GIS (Geographic Information System) software. This presentation is an introduction to the hardware, software, data, and expertise required to map genealogical data. Data about cemeteries in Jackson County, North Carolina, USA are used to demonstrate how genealogical data can be mapped using GIS software.

# What Are Geographic Information Systems

GIS is a computer-based tool for mapping and analyzing features and events on earth. GIS technology integrates common database operations, such as query and statistical analysis, with maps. GIS is a tool used by individuals and organizations, such as schools, governments, and businesses, seeking innovative ways to solve problems and increase the quality of decision making. Click any of the multimedia titles below for more information.

[What Is GIS?](#)

[GIS.com](#)

[GIS Day](#)

HOME Select a category for more information.

© HELP EXIT GIS Software Industries About ESRI About GIS

So what is GIS? In very simple terms it is a computer mapping program plus a database management system that allows you to analyze data and map the results. It is what cartographers have done for thousands of years: draw maps, interpret thematic data, and add the thematic data as layers to the maps. The only difference between the old and the new is that in the past all of the work was done manually and now computers do much of the work. But the concepts and the knowledge required to draw maps have not changed.

The data in a GIS system is organized into layers depending on what type of feature the data represents. There are basically three types of features: points, lines, and polygons. Point features include such things as houses, springs, oil wells, etc. Line features can be rivers, roads, political boundaries, etc. Polygons represent geographic areas such as continents, oceans, etc. as well as political entities such as countries, states, provinces, etc. A layer should include only one kind of feature within a category, i.e. houses and churches should not be in one layer even though they are both point features.

What makes a GIS system so flexible is that you can edit a layer without affecting any other layer. You can also statistically manipulate the attribute data associated with a layer. For example, you can group data into a number of classes and these classes can be determined by using different

classification systems. But you are not restricted to analyzing one layer at a time. You can also analyze data in one layer in terms of another layer. For example, you could determine which bus routes (line layer) are within two blocks of a shopping center (point layer).

## **Components of a GIS System**

- 1. Hardware**
- 2. Software**
- 3. Data**
  - a. Spatial
  - b. Attribute

In order to do GIS analysis, you will need a GIS system. The components of this system include a computer, GIS software, and spatial and attribute data. If you plan to print out any of the maps, a color printer designed to produce fairly high resolution images should be added to the list. Since maps are a form of graphic representation, the hardware requirements are similar to that for graphic programs.

# Hardware

- **Operating System**
  - Windows NT 4.0 or
  - Windows 2000 or
  - Windows XP (Home or Professional Edition)
- **Intel Pentium Processor**
  - 1.0 GHz recommended or higher
- **Memory/RAM**
  - 512 MB minimum
  - 1 GB recommended

For GIS software to work efficiently, you will need a high-end (powerful) computer. In a graphics program every pixel must be accounted for. This creates large files and large files require fast processors and large memory capacity. Thus a 1.0 GHz Intel Pentium Processor or higher and a minimum of 512 MB of RAM (preferably 1 GB) is recommended. Lots of RAM makes things happen more quickly. The data files from which the maps are generated as well as the end product tend to be large, so large capacity storage devices are also recommended. Most GIS software runs on a Windows operating system. If you are acquiring a new system, Windows XP (home or professional edition) would be the operating system of choice. However, the software will also run Windows NT 4.0 or Windows 2000. If you have the resources to acquire a more powerful computer than the minimum requirements, it would be wise to do so. New versions of GIS software tend to require more and more power in terms of processors and RAM.

# Software

- **GIS software companies:**
  - ESRI (Environmental Systems Research Institute)
  - MapInfo
  - Intergraph
  - Autodesk (product of AutoCAD)
  - Microsoft MapPoint

There are many GIS software programs available. The software you choose, will depend on what your needs are. The cost varies from free to very expensive and the capabilities of the software vary from very basic to very sophisticated. There are also numerous websites that offer GIS capabilities. These usually have limited functions and you are limited to the data that is available at that website.

Listed above are some of the major producers of GIS software. ESRI (Environmental Systems Research Institute) is usually considered to be the leader in the field and the standard setter. They offer three basic versions of their ArcGIS software. ArcView is the basic version and offers enough functionality to meet most needs. ArcEdit, in addition to the ArcView functionality, allows you to edit the spatial data. ArcInfo is the high end product and includes all of the functionalities of ArcView and ArcEdit plus many more. ArcInfo is intended for those users who do extensive analysis and modeling. The cost, of course, also increases as functionality increases. Thus ArcView is the least expensive and ArcInfo the most expensive. ESRI also offers a number of extensions that run on top of the ArcGIS products and extend their capabilities.

The MapInfo product, MapInfo Professional, is very similar to ArcView but less sophisticated. It is owned by SPSS (Statistical Program for the Social Sciences). Cost is about the same as for ArcView.

Intergraph originally produced only CAD software (Computer Aided Design). It now also offers a GIS product called GeoMedia.

AutoDesk is a GIS software package produced by Autocad, a standard setter in CAD (Computer Aided Design) software.

Microsoft has also entered the GIS world with a product called MapPoint. It is aimed at the casual user and the cost is somewhat less than that of the other GIS software packages. Microsoft also offers an online service called MapPoint.Net.

Many of these companies offer educational discounts to qualifying institutions. In some instances this reduces the cost of the software by as much as 80%.

## Data

- **There are two types of data:**
  - Spatial Data
  - Attribute Data
- **Data are available from:**
  - Government agencies
  - Commercial sources
  - Colleagues, etc.
- **Site specific data often not available**

A GIS system requires data to create maps. This data, both spatial and attribute, is stored in tables. The GIS software reads these tables and transforms them into layers that together create the map you see on the screen. Data is the most important component of a GIS system. Without it you cannot do any analysis. Unfortunately data is also the most expensive and most difficult to acquire. To complicate matters even more, you need two types of data when doing GIS analysis – spatial and attribute. Spatial data is what the software uses to create the map. There are three types of spatial data: point, line, and polygon data. So if you want to show the cities, roads, and political subdivisions of a country, you would need all three types of spatial data. Attribute data is information about the spatial entities and must be for the same geographic unit that is represented on the map. Thus if you have a map that shows provinces, then the attribute data would also have to be broken down by those same provinces. Frequently spatial data includes attribute data, but this is not always the case.

High-end GIS software, such as ArcInfo, allows you to edit existing spatial data as well as create your own spatial data. Low-end GIS software has limited capabilities to edit or create spatial data. However, most GIS software allows you to join existing tables of attribute data to tables of spatial data. The only requirement is that the two tables have one field in common. For example, if you have a table of attribute data about cities, you can join it to a table of spatial data if both have a field that gives the name of the city.

Data can be obtained from a variety of sources including government agencies, commercial companies, and colleagues or other individuals who may have already created the data you need and are willing to share it with others. In some countries, such as the United States, geospatial data from government agencies is usually free or available for a minimal fee. Government agencies in many countries, however, charge for their data. There are also many commercial companies that create or repackage data. It is often advisable to purchase data from these companies, even if it is a little more costly. Free data often contains errors and it is more costly to correct these errors than to purchase data that has already been edited for errors. However, if you need very specific data for a very specific area, it frequently is not available from any source. Your only recourse in such a situation is to create it yourself. This can be both time consuming and costly.

## **Why use GIS for Genealogical Research?**

- **Display the spatial distribution of data**
- **Detect patterns that are not easily apparent from tables**
- **Analyze data and map the results to identify new patterns**

So why use GIS in genealogical research? It allows you to map the spatial component of genealogical data. Mapping this information reveals spatial patterns that are often overlooked or not obvious in a table. The saying that a picture is worth a thousand words certainly applies to maps. Maps, after all, are a “picture” of the earth. What makes GIS even more useful in

genealogical research is that you can take existing data, analyze it in a variety of ways, map the results, and perhaps identify new patterns that were not readily apparent in the original data.

# **The Cemeteries of Of Jackson County, North Carolina USA**

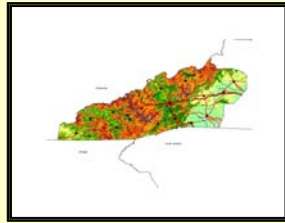
A Demo

Rather than talk more about the capabilities of GIS software and how it can be used by genealogist, I would like to demonstrate how GIS can be used in genealogical research.

Data for this demonstration were taken from the publication The Cemeteries of Jackson County, North Carolina compiled by the Jackson County Genealogical Society. This publication contains 550 pages and represents the effort of a large number of people. The data used for this demonstration represents only a small fraction of the wealth of information included in the publication.



## Where is North Carolina?



North Carolina is located in the southeastern part of the United States, commonly known as the “South.” Jackson County is located in the western part of North Carolina. It is a mountainous region and until recently was fairly isolated not only from the rest of the state of North Carolina but also from the United States. The area was settled primarily by Scotch-Irish settlers in the early nineteenth century. Because of the remoteness of the area, many of the customs of these early settlers have survived. One of the unique characteristics of western North Carolina as well as other areas of the South is the number of cemeteries that are scattered throughout the countryside. Many churches maintain their own cemeteries, local communities have their own cemeteries, and then there are family cemeteries as well as municipal cemeteries. In Jackson County (population about 36,000), for example, there are 231 known cemeteries and most of them are located on hillsides and in remote areas.

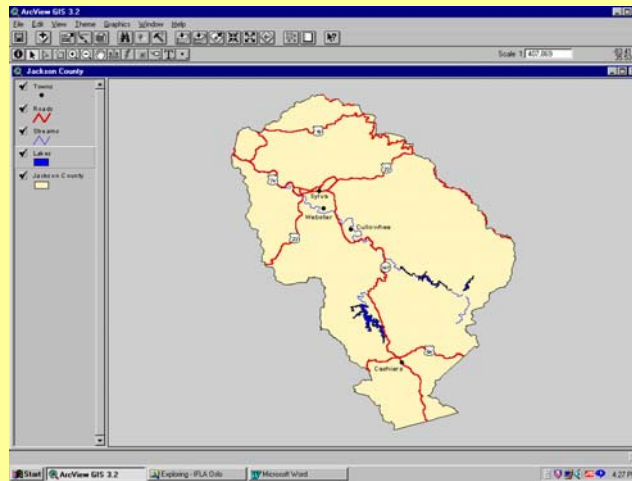
## The North Carolina Mountains in Pictures



As you can see the mountains are heavily forested due a warm climate and abundant rainfall. Unless the cemeteries are maintained, they are quickly overgrown with native vegetation and forgotten. Yet these cemeteries are an important link to the past. Record keeping in the early days was not always what it should have been and often the gravestones are the only source of information. With this in mind, let us see how we can map some of the information gathered from these cemeteries.

# Getting Started

## Step 1: Create a Map of Jackson County



The first step is to create a base map that can be used to add genealogical data. In this case it was possible to extract the data to create the county map from a larger data set. Although the map appears to be one layer, in reality it consists of five layers with each layer representing a different type of feature. The layers are listed on the left side of the illustration and in the order in which they are arranged on the map. The polygon layer(s) (layers representing areas) are always at the bottom, followed by layers for line features, and the top layers are always point features. Thus the top layer shows the towns. This is followed by two line layers, one for roads and one for streams. The bottom two layers represent the polygon features of reservoirs and the county outline. Note that even though roads and streams are both line features and reservoirs and the county outline are both polygon features, they are in separate layers. This is because you can edit each layer without affecting the other layers, but within a layer all the features are affected by any changes made.

## Step 2: Create a Table of Jackson County Cemeteries

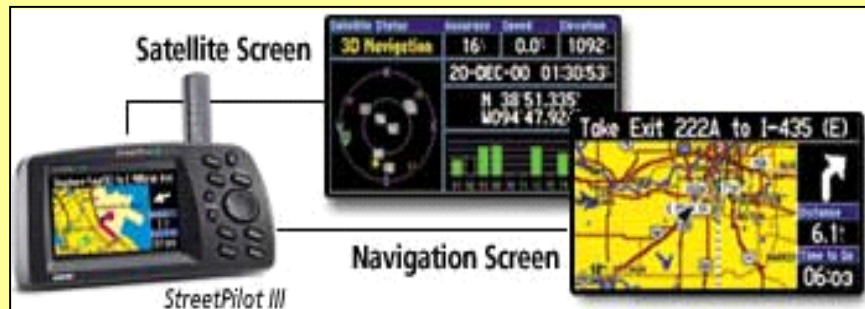
Name	Address	City	Latitude	Longitude	County	
Point - Babson Grove	Fish Mountain	Tuckahoe	39.1906	83.0342	35, 31.8333	-83.16188
Point - Elm Knob	Dicks Creek	Whitaker	35.2144	83.1931	35, 28.855	-83.26388
Point - Elm Ridge		Big Ridge	35.1143	83.0506	35, 1.8666	-83.05500
Point - Flanston/Higley	Nassau	Salva North	35.2329	83.0809	35, 39.113	-83.15138
Point - Gooden (Baker's Creek)		Whitaker	35.2525	83.0112	35, 1.9827	-83.30333
Point - Groves		Tuckahoe	39.1627	83.0241	35, 27.416	-83.04472
Point - Groves (Baker's Creek)		Tuckahoe	39.1816	83.0513	35, 30.444	-83.02960
Point - Groves (Lester's Creek)		Tuckahoe	39.1653	83.0416	35, 29.39	-83.07111
Point - Hangerman		Whitaker	35.2413	83.1931	35, 40.263	-83.25217
Point - Ledge		Genova Creek	35.2033	83.1534	35, 14.950	-83.25444
Point - Lefkey		Tuckahoe	39.1906	83.0141	35, 25.166	-83.02895
Point - Clayton		Salva North	35.2427	83.0134	35, 40.790	-83.15944
Point - Leonard (Lanes)		Tuckahoe	39.1796	83.0461	35, 29.688	-83.11333
Point - Leonard (Sage)		Tuckahoe	39.1712	83.0293	35, 28.666	-83.06811
Point - Lion Creek		Salva South	39.1763	83.1123	35, 29.665	-83.18932
Point - Crawford		Headwood	35.2562	83.0646	35, 41.712	-83.11277
Point - Calhoun's Baptist Church		Salva South	39.1893	83.0462	35, 31.388	-83.17923
Point - Calhoun's Methodist Church		Salva South	39.1834	83.0468	35, 30.844	-83.18880
Point - Oak (Oak Creek)		Salva South	35.2349	83.1251	35, 30.000	-83.23883
Point - Oak (Oak Creek)		Salva South	35.2357	83.1136	35, 30.816	-83.13333
Point - Double Springs		Genova	39.1243	83.0993	35, 21.81	-83.19385
Point - East Fork		Genova Creek	39.1838	83.1522	35, 31.995	-83.25811
Point - East		Genova	39.1796	83.1025	35, 1.8666	-83.17383
Point - Fairview Memorial Gardens		Salva South	35.2162	83.1157	35, 35.995	-83.18916
Point - Franklin (Long Branch)		Whitaker	35.2228	83.1543	35, 37.722	-83.26184
Point - Garter (Garter Hill)		Genova Creek	35.217	83.0305	35, 36.627	-83.24472
Point - Garter Hill's Chapel		Whitaker	35.2654	83.1769	35, 44.833	-83.28972
Point - Hartsburg		Genova Hill	39.0953	83.0799	35, 14.472	-83.12760
Point - Henry		Salva North	35.2418	83.0605	35, 40.900	-83.13472
Point - Hensley		Whitaker	35.2480	83.0195	35, 40.000	-83.14416
Point - Hope		Salva South	35.2341	83.0837	35, 37.883	-83.14461
Point - Hyatt		Whitaker	35.2654	83.1769	35, 45.833	-83.28972
Point - Jones		Headwood	35.2513	83.0524	35, 42.627	-83.00688
Point - Kester (Fisher)		Salva North	35.2342	83.1251	35, 37.833	-83.23883

The next step is to add a layer to the map that shows the location of the cemeteries. However, spatial data for the cemeteries does not exist and thus it was necessary to collect the necessary data from various sources. In order to show the cemeteries on the map, the exact location of the cemeteries has to be determined. This is done by defining the location using a coordinate system that the GIS software understands. The most common coordinate systems used include latitude/longitude, the UTM (Universal Transverse Mercator) coordinates, State Plane Coordinates (for the US), and a variety of other coordinate systems.

Once the coordinates have been determined, they need to be put in a table. This table should show the latitude and longitude readings (or equivalent coordinate system readings) in separate fields. It should also have a field that identifies the name of the cemetery. Other fields can be added with additional information. Once this information is in the proper format, the GIS software can convert the coordinates into a form that the software can then use to plot each cemetery on the map.

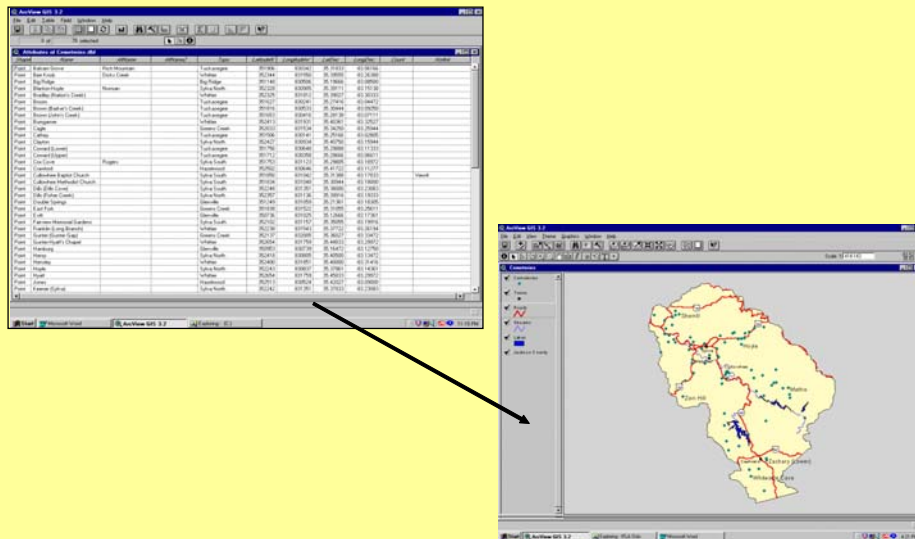
However, identifying the coordinates for each cemetery can be just as difficult as finding existing spatial data. For example, it was possible to identify the coordinates for only 78 of the 231 cemeteries by using the US Geological Survey's Geographic Names Information System (GNIS). The coordinates for the remaining 153 cemeteries (about two-thirds) have to be determined in some other way.

**Step 3: Use GPS Receiver, if Necessary, to Determine Coordinates**



The modern way to determine coordinates is to use a GPS (Global Positioning System) receiver. With this little marvel of modern technology you can go to the location in question, determine the coordinates for the feature, and then manually enter them into a table, or, with proper software, load the coordinates directly into the computer. The coordinate system used, however, must match that of the map. A not so high-tech method is to use a roamer and determine the coordinates using a paper map. A roamer is a transparent template on which the degrees of latitude/longitude, UTM units of meters, etc. are broken down into smaller units. This allows one to determine the coordinates of a feature fairly precisely.

#### Step 4: Convert the Cemetery Table to a Layer on the Map



Once all of the coordinates for the cemeteries are identified and attribute data, if any, about the features are added to the table, the table can be converted to a layer that shows where the cemeteries are located. This layer can then be added to the existing map.

However, rather than just show where the cemeteries are located, it would also be useful to know something about the graves that are located in these cemeteries. Since there are many graves in a cemetery, information about the graves cannot be added directly to the table of cemeteries. But it is possible to create a new table for the graves and add the data in summary form to the cemeteries table.

## Step 5: Create a Table of the Graves

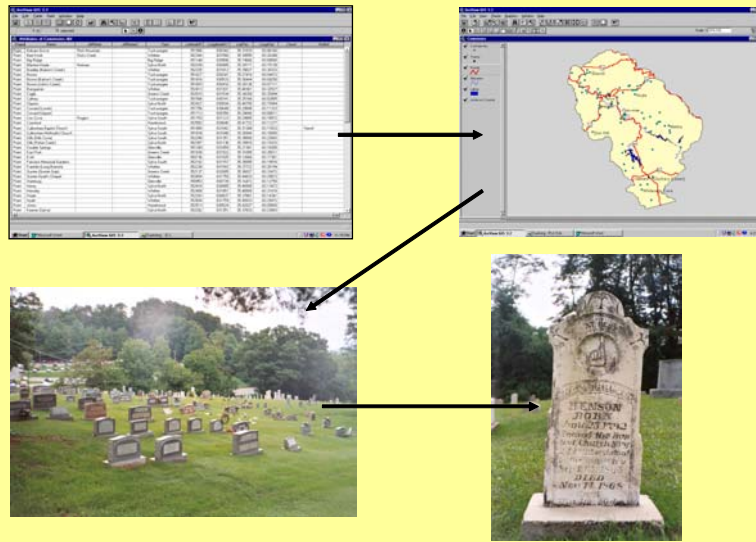
Surname	First Name	Middle Name	Maiden Name	Birth	Death	Cemetery	Age at Death
Cogbill	Lee	Ray		1911011	1956008	Adde Community	1
Cogbill	Walter	D.		1914019	1982208	Adde Community	1
Cogbill	John	A.		1916046	1946026	Adde Community	1
Volord	Elta	N.		1916021		Batum Grove	1
Volord	Emp	Valene		1942011	1956016	Batum Grove	1
Volord	Howard	C.		1907009	1910121	Batum Grove	1
Volord	Emma	C.		1897014	1963000	Bea Knobs	1
Volord	Harold	Wf		1926019	1969010	Bea Knobs	1
Volord	Earl			1926022	1965122	Bea Knobs	1
Volord	Juana	M.		1917018	1969020	Bea Knobs	2
Volord	John	D.		1914023	1991005	Big Ridge	2
Volord	Robert	Ray		1918000	1979000	Big Ridge	1
Volord	Debra	A.		1956022	1990111	BlantonHigle	1
Volord	Judith			1900211	1978028	BlantonHigle	2
Volord	William	B.		1847126	1930016	BlantonHigle	1
Volord	John			1907012		Bradley (Robert's Creek)	1
Volord	Nancy			1906019		Bradley (Robert's Creek)	1
Volord	Claude	H.		1916044	1976017	Bradley (Robert's Creek)	1
Volord	Glad	Denny		1912229	1947010	Broom	1
Volord	W			1904002		Broom	1
Volord	Gary			1963020		Broom	1
Volord	Elaine	J		1925026		Broom (Robert's Creek)	1
Volord	Ben	D.		1906011	1976001	Broom (Robert's Creek)	1
Volord	W			1923017	19811016	Broom (Robert's Creek)	1
Volord	Lee	L. (B)		1896016	1964013	Broom (John's Creek)	1
Volord	Edgar			1896016	1964013	Broom (John's Creek)	1
Volord	Robert	D.		1947021	1964024	Broom (John's Creek)	1
Volord	Joseph	A.		1867010	1926026	Bungarum	1
Volord	Fernis	A.		1878000	1951000	Bungarum	1
Volord	Geniel	C.		1927022	1949014	Bungarum	1

If you want to use the information from the table of graves in conjunction with the cemeteries layer, the table must include a field for the cemetery where the grave is located. It is this field that will permit you to join the data about the graves to the cemeteries table. In addition to the field for the cemetery, additional information can be added in separate fields. Obvious information would include surname, first name, middle name, maiden name, birth date, death date, and any other information that is deemed important.

Now that the basic data is in place and in a format that GIS software can understand, it is possible to analyze/manipulate the data in a variety of ways and then map the derived data. The following are a few examples.

# Ways to Analyze & Map the Data

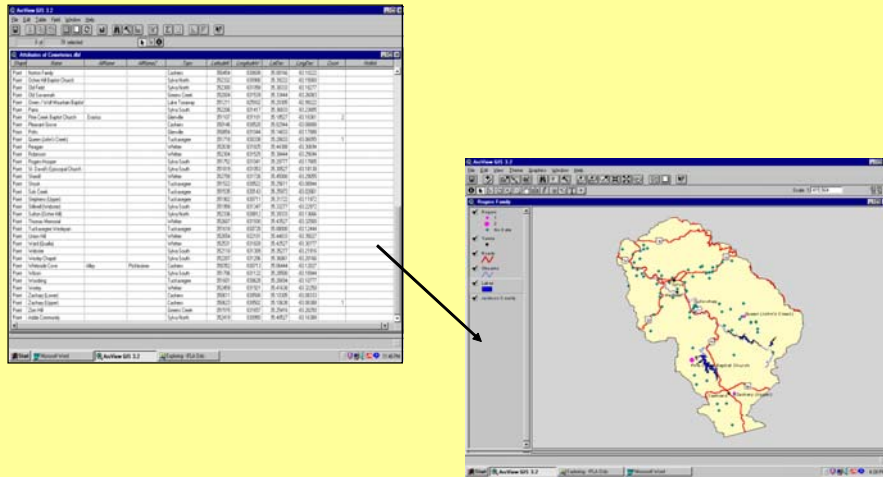
## Scenario 1: Create a Link From a Cemetery to a Picture or Other Document



This first example shows how various documents can be linked together. If a field is added to the cemeteries table that shows the location of another document, it is possible to create a link to that document. In this case it is to a picture of the cemetery. If the picture were overlaid with a grid, a second link could be created to show another document (in this case a picture of the tombstone). The linked document does not have to be a picture. It could be a table, a graph, a legal document, etc.



## Scenario 2: Identify Cemeteries by Surname



This second example identifies cemeteries where individuals with the same surname are buried and the number of graves. To accomplish this all the records with the same surname are selected in the table of graves. This list is then summarized by the name of the cemetery and a table that shows how many graves with the same surname appear in each cemetery is automatically generated. This summary table can then be joined to the cemeteries table and the data mapped. On this particular map a graduated symbol is used to display the distribution.



## Summary

- **GIS is a tool that can be used to visualize genealogical data**
- **To do GIS analysis you need a computer, GIS software, spatial data, and attribute data**
- **Spatial and attribute data is the most costly component**
- **Data can be analyzed in a variety of ways**
- **The resulting data can then be mapped**
- **Maps often reveal patterns that are not evident in table format**
- **The down side is that there is a steep learning curve associated with learning GIS**

These are just some examples of how GIS can be used to visualize genealogical data. Some simple analyses were performed on the original data to create new data. This newly created data was then mapped to perhaps reveal patterns that were not obvious from the tables.

To do GIS analysis requires a high-end (powerful) computer, GIS software to meet your needs, and data. The data are usually the most costly component of the system whether it is purchased or collected by the user. The results of GIS analysis, however, can be very useful and rewarding. Unfortunately the learning curve associated with GIS is very steep, but once mastered, GIS is a very useful tool in genealogical research.

## Bibliography

The Cemeteries of Jackson County, North Carolina. Cullowhee: The Jackson County Genealogical Society, Inc., 1998.