3 Topics for Sociospatial Research

Chapter Description

In this chapter, we explore how incorporating a spatial context into your investigation can enhance your analysis by providing additional insights and information not previously considered. We first explore approaches for defining project goals, and, based on the defined goals, we explore whether GIS analysis techniques enhance the outcome. The chapter includes sections on developing your project goals and approach, including relevant examples that explore the potential of spatial analysis as part of social science research. We also include suggestions for the application of these approaches when locating or collecting data for your own research. These examples illustrate how common data types can be tied to the spatial realm of the GIS.

Chapter Objectives

- Introduce GIS as a tool for the integration and analysis of social science data.
- Offer a series of considerations to use in defining the role GIS will play in your research.
- Provide examples of research that would benefit from a spatial perspective.

After reading this chapter you should be able to perform the following tasks:

- Select, organize, and begin to seek out relationships in a variety of data.
- Consider a variety of potential factors related to social data that might help to explain observed differences between locations.
- Define and begin to conceptualize goals for a GIS project, including the spatial extent of your study area, the data needed, and the approaches to analysis.
- Begin to plan approaches to data collection, including the value of taking a computer with you into the field.

Introduction

Although GIS does not have a long history in the social sciences, its value in social science research is beginning to be recognized. GIS is such a powerful tool for the social scientist because it allows for the integration and comparison of contextual data from a social as well as environmental or physical standpoint. Almost all data that the social scientist examines have an associated geographic point of location. The ability to identify where differences or similarities, correlations and interactions exist is of major import to the social science researcher. Additionally, GIS can facilitate the incorporation of both qualitative and quantitative variables into one's study. This is discussed in more detail in Chapters 4 and 5.

What Value Does GIS Present in Social Science Research?

Analysis may, on the surface, seem an obvious reason for getting involved with a technology such as GIS. Indeed, the purpose of this book is to present a variety of ways that GIS can be effectively incorporated into your own research. Surprisingly, many people who use GIS use very few, if any, of the analysis tools offered in any of the commonly available software packages. Instead, the GIS is viewed simply as a tool for designing simple maps. Certainly, a well-designed map of your data can communicate a tremendous amount of information, and this is one component of the GIS. A good book on cartographic design will prove a valuable resource for map design. We address map design briefly in Chapter 2.

The real value of GIS to the researcher comes in the analytical capabilities added when incorporating the spatial component GIS provides. This is a major focus of this book, particularly in Chapter 9. Before discussing analysis, it is important to provide a basic overview of GIS organization and structure as well as data formats that can be most easily incorporated. One especially appealing benefit is that most, if not all, of the data already being collected by social science researchers can be easily integrated into a GIS, and, with a little extra up-front planning, all data can be prepared in a GIS-friendly format. Data are the focus of Chapter 2; data conceptualization in a spatial context is further discussed in Chapter 4. Chapter 6 focuses on the collection of spatial data in your research process and the location or entry of data that will be GIS friendly when it comes time to incorporate the data into your GIS analysis.

Exploring and Integrating Information

Studies taking into account a variety of contextual and individual-level variables would benefit from using GIS. Why? GIS allow researchers to

establish a context for their individual-level data. For example, if you were to survey a group of people living in a particular state, a GIS would allow for the specific connection of each individual's responses to a geographic location, such as a particular community within that state. These individuals do not exist in a vacuum; they live in particular places. Although it may be perfectly desirable to understand attitudes of respondents statewide, there is potentially much to be gained by understanding differences in different regions of the state (e.g., urban vs. rural counties). Similarly, any study can be examined at a variety of spatial scales, and the unit of analysis appropriate for one study may be entirely different from that of another study.

Capturing the additional location information is useful when conducting more targeted analyses, for understanding the social context pertaining to a particular geographic area, or for understanding characteristics of an area that might influence the social situation. A GIS facilitates the integration of spatial and nonspatial data because the researcher identifies which variables (environmental, physical, individual) to incorporate into the analysis. As long as the variables in your study can be tied to a geographic location, they can be incorporated into a GIS. Later in the book we also discuss how GIS can be used to analyze data that are not linked to a geographic location.

A GIS is perfect for integrating physical, environmental, and social data in one unified analysis environment. One benefit comes via the visualization of the data as your analysis proceeds; that is, you can see how things appear to relate in space by viewing them on a map before moving forward with a formal analysis to explore the statistical significance of the relationship. A GIS may also allow the researcher to examine how previously unconsidered geographic variables integrate into the system, making the GIS an excellent data exploration tool. Because a GIS can integrate data from a variety of sources, it allows the researcher to develop a holistic view of the many different contextual variables that may be important to addressing a particular social issue or research question.

For instance, let's say that you are studying a community where homelessness is considered a major problem. Residents of the local community complain to city officials about the constant panhandling that occurs around the community as the homeless try to make ends meet. City officials could commission a study of the migration pattern of the homeless by conducting interviews to determine where geographically these individuals are spending their time, where they are sleeping, and what sorts of needs they have (e.g., public bathrooms and food distribution). Using a GIS to map the location and needs of the homeless community and the areas that currently serve some of their needs, gaps, or shortfalls in services might be identified. In concert with a variety of social, environmental, and physical data, in combination with public input, the GIS could be used to identify optimal and willing locations to provide additional services to the homeless community and to more effectively help them to get back on their feet.

Seeking Relationships

A GIS facilitates the examination of relationships between different types of data. A GIS allows for the connection of both qualitative and quantitative data in a geographic context. An example of the incorporation of qualitative data might be the oral histories or stories told by the elders of a particular tribe or community about the importance of nature to their local community. This data could be linked to specific geographic locations within the community, for example, locations of traditional hunting and fishing grounds, ceremonial sites, or significant landmarks of cultural or personal significance. The GIS would facilitate the exploration of relationships between individuals' geographic locations within the community and the importance the individuals place on nature. For instance, elders who live closer to the edges of the community and in closer proximity to the surrounding natural forest may emphasize more the importance of nature than those who live in the more urban or central part of the community.

Another interesting function of the GIS might be to identify areas consistently mentioned by multiple respondents interviewed during the study to identify the common, shared history and understanding of the community. Using a GIS facilitates the visualization of such relationships. Correlations that may not be readily apparent in a traditional analysis may jump out when viewed spatially on a map. Furthermore, if the GIS is available in the field during the interview process, respondents could point out locations on the map, further enhancing the process of data collection and analysis. Using maps, photos, or other visual information can be especially useful in situations where language is a barrier to data collection.

Thinking Critically

The ability to think critically is a skill needed not only in academia but also in the corporate world. But what does it mean to think critically? Thinking critically means not relying solely on other people's explanations or rationalizations for phenomena. Critical thinking is a major skill that is stressed in many universities throughout the country. Using a GIS facilitates critical thinking because it allows for the integration of previously unconsidered geographically based variables into one's study. This integration contributes to critical thinking because it helps the researcher to move beyond the consideration of traditional variables or explanations for observed changes and explanations in one's study. It facilitates the inclusion of both qualitative and quantitative analyses into a study.

Qualitative analysis is defined as "the nonnumerical examination and interpretation of observations for the purpose of discovering underlying meanings and patterns of relationships" (Babbie, 2003, p. G6), whereas quantitative analysis is defined as "the numerical representation and

manipulation of observations for the purpose of describing and explaining the phenomena that those observations reflect" (Babbie, 2003, p. G6).

Acknowledging Differences

A GIS can be very useful for examining differences that exist between variables in your study. A GIS could help a researcher to identify differences between groups, places, organizations, and resources. The list is almost endless. As long as the data can be linked to a geographic point, they can be examined using GIS. In Figure 3.1 we considered the possibility that differences across the state of Illinois might have something to do with urban versus rural settings as defined by population. Further analysis might explore relationships to other variables that are socioeconomic (age, income, education), environmental (proximity to Lake Michigan, inland lakes or rivers, forest, parks), or physical (roads, telecommunications, housing). We might find that a variety of variables interrelate in space in a meaningful way and are worthy of further exploration.

For example, a GIS can be used to compare variable attributes at different points in time (assuming that you have data for these different time

Rockford: 115,000
Chicago: 2.9 million

State: 12.4 million
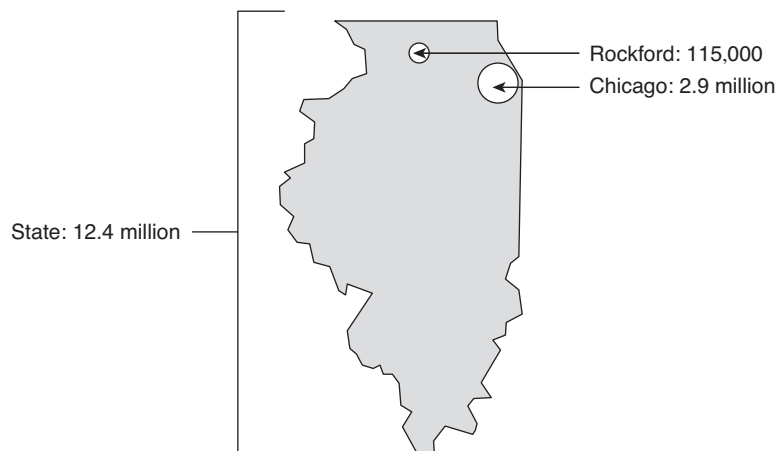
Figure 3.1   An analysis of Illinois could be easily biased by the large urban population in the greater Chicago area if the spatial component of sample data is not recorded for inclusion in a spatial analysis. In a state such as Illinois, Chicago is the single largest city by far. Chicago has nearly 2.9 million people and is about 23% of the population of the entire state. The next largest city, Rockford, has a population just over 150,000, about 5% of Chicago's population (U.S. Census, 2000). An analysis of Illinois that did not incorporate spatial information could be biased by the large urban population in the greater Chicago area representing much of the northeastern part of the state. The geographic location of respondents in a survey in Illinois might significantly influence the kind of responses obtained and could be overlooked if the spatial component is not recorded for inclusion in a spatial analysis.

periods). An overlay of variables for two or more points in time can be used to analyze actual social and environmental changes that have occurred for a particular study site over a period of time. This type of longitudinal study is sometimes termed a temporal analysis. If we were to consider just the example of Chicago with its 2.9 million people in 2000, one might ask how Chicago changed in the 10 years from 1990 to 2000. The census shows an increase of 3%, or about 112,000 people, enough people to match the size of the fifth largest city in the state! Granted Chicago is a large city, so one might reasonably ask, "Where in the city, which neighborhoods, did these all these new people move into?" With the data as stated here, you would have no idea, but with a little additional information (census block data), you could identify those areas of the city that experienced the most growth, those that stayed the same, and those with a decreased population over those years.

Census-based analyses of this sort are convenient because the data are already collected in a manner appropriate to these questions and are therefore an excellent source for many GIS-based analyses of social science data. It can be slightly more difficult to assemble data necessary to investigate other questions where the data are not quite so easily acquired from one source, but the process would be similar.

Finding Common Ground

A GIS can assist in the identification of common ground between social and environmental variables. In other words, it allows for discovering where different groups, clusters, towns, and corporations share common issues. GIS technology allows for examining the geographic intersection between different environmental and social features.

Additionally, a GIS allows for the visual portrayal of data. Having this ability is essential for communications between people of different backgrounds and can help groups to realize that they share things in common. Identifying common ground is often the first step toward building a consensus out of conflicts that often occur in planning. For instance, let's say that you have a consortium of service organizations that meets with the goal of streamlining the organizations' contributions to the community. Using a GIS enables the entire collection of service organizations to break up services to the community by the area of town that is closet to their club (Figure 3.2). So, for instance, the Southside Rotary Club would put its resources into improving playgrounds on the south side of town, while the Northeast Lions Club would focus on improving services for senior citizens on the northeastern side of town. This would be an approach based on separating the city into areas in nearest proximity to each organization, also known as proximity polygons (discussed further in Chapter 9).
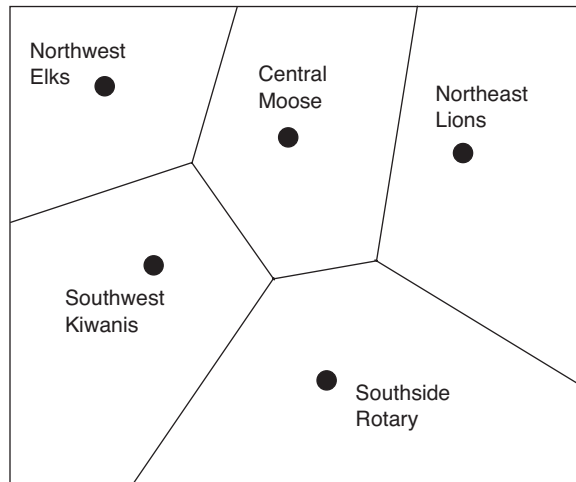
Figure 3.2   Proximity polygons delineate the areas nearest to each of the service organizations in the community. Everything within the polygon surrounding a particular point is closer to that point than to any other point on the map.

Determining Project Goals

Any person interested in conducting research ultimately begins determining the main idea, or purpose, of one's research. For the social sciences, the possibilities are endless, especially when it comes to incorporating GIS into the research plan. There are many different factors that can affect the purpose of your research. Some of these include your own personal interests and the interests of the sponsoring organization or agency. Applied research projects may have more than one single purpose, setting out to answer a series of questions or to work toward a solution to specific social problems.

Anytime you engage in a research project, you need a clear conceptualization of what the research goals are. A surprising number of people carrying out an analysis cannot easily answer the simple question, what is the purpose of this study? Having a clear project goal is very important because, beyond being the reason for why you are engaging in research in the first place, the goal directly informs the data you need to locate, collect, and prepare for the eventual analysis. A project may have more than one goal and different participants in and funders of the project may have different goals. If this is the case, it is useful to develop a set of overarching project objectives and a list of needed data prior to beginning the data collection process. One of the most troublesome things that can occur in carrying out an analysis is realizing after the fact that an essential and easy to acquire piece of data was not collected as part of the original design. Going back to fill in missing data after the fact, assuming it is possible at all, can be time-consuming,

expensive, and of course embarrassing. In the GIS realm, you will also need to determine how you will encode the data in a GIS compatible format that simultaneously works within the GIS and effectively meets the requirements of your study. Chapter 2 addresses these issues in more detail.

Some questions you might ask yourself in developing goals for your study include the following:

- Why are you engaging in this study?
- What sorts of relationships do you want to examine as a part of your study?
- What causes or organizations might your study benefit?
- Who is funding your study?
- Are you studying patterns of human behavior?

If you are able to answer these questions clearly and up front, you will have a much easier time defining the specific goals of the project and thus the data and variables you will require in achieving your objectives.

Guiding Questions

To discover whether your proposed study could be enhanced by integrating geographic information you can begin by answering a few questions.

The following sections describe four different sets of questions for researchers to consider in determining how a GIS can be integrated into their study. The sections include questions about concept, questions about data, questions about location, and questions for analysis.

Questions About Concept

The questions in this section focus on the ideological structure of your study. These questions are designed to be asked early in the research process, usually before you begin implementing your research methodology. When answering these questions, some background research may be necessary; this research should be done before beginning the actual project and going out into the field. You may be able to conduct much of your background research through a literature review and an Internet search or through talks with local experts and other individuals familiar with the study site to gain a good understanding of the existing relationships. However, your contact with people in your study site will depend on whether or not you have established social ties there, how accessible or remote your study site is, and the availability and access to technology (phones, the Internet, etc.) at the site. Sometimes a small pilot study is useful to ensure your approach is acquiring the necessary data and is not

confusing to those doing the collection or to those responding to questions. Pilot testing of the analysis can also be helpful to ensure that the data can be analyzed as expected. Planning and carrying out the GIS analysis is discussed in detail in Chapters 4 and 9.

Of course, it is important to collect the proper information from each location, respondent, or unit of analysis in your study. This information is commonly referred to as variables. What is a variable? A variable can be defined as a domain of attributes that relate to a particular concept. For example, the variable Gender might be defined as the three attributes: male, female, or other. So what is an attribute? An attribute is a characteristic that is associated with a particular object or person. The following questions might help you in defining the concepts to be used in your study:

- What is your research question?
- What is your main dependent variable?
- What are your independent variables?
- What are the main hypotheses of your study?
- Are there any geographic features that you have already identified as variables in your study? If so, what are they? Are they man-made or natural features?
- Could any of the geographic features contained in your study site potentially affect the issue that you are studying?

Questions About Data

The questions in this section relate to the data collection portion of your project. Once again, it would benefit the researcher to find out answers to these questions prior to going into the field. This information should be available from local government agencies (local, state, county, federal) that have responsibility for the area, including your study site. In addition, contacting private businesses or local interest groups may turn up even more of the information you require. As was mentioned earlier, one of the keys to the successful use of GIS is being able to be a creative thinker and seeking creative avenues to find the data that you might need for your study. Although you may want to simply search the Internet for agency and organizational Web sites to get all the data you need, this process will rarely provide anything close to a complete list of the data that are actually available. More likely, you will need to make direct contact with the appropriate people at each agency or organization to find out answers to these questions. Because the collection and development of quality data sets are time-consuming and expensive processes, some people will be reluctant to provide their data to you. Therefore, it is often worthwhile to have something to offer in exchange for the data. Data trades, access to related data to be developed during the study, or access to the final report results are all valuable offers you can

make to those who assist you. When considering the use of GIS for any particular study, some items to consider include the following:

- Do GIS data exist in any form for your study site?
- If not, what sort of data exists that might be GIS compatible?
- Do you have access to this data and, if so, under what conditions?
- What was the original purpose for which the data were collected?
- What variable attributes are included in the data set that you are interested in examining?
- How old are the data? Were the data collected at only one time, or are the data longitudinal?
- What sort of metadata exists for these data?
- Has any type of accuracy assessment been completed, or have the data been ground truthed?

Assuming that you can find and access data appropriate to your study, the time spent locating useful data, especially if it is already in a GIS format, will be well worth the effort. Even finding out that none of the required data exists or is available is still important information in planning your study. Knowing what you will need to collect and input data yourself is important in adequate budgeting of both time and money for your study. Considerations of data quality are addressed in Chapters 2 and 8.

Questions About Location

This section contains a series of questions that should be asked about your geographic location. These questions may appear basic, but they help you in gathering information that is useful for integrating a GIS into your project. These questions will help you in the process of identifying important characteristics of your study location. Answering these questions beforehand assists you in your data collection process because it helps you establish those aspects of your location that are important and determine which type of data variables you require. It is important you answer these questions before data collection so that you can arrange the categories and measurement devices for your project. These issues are discussed in detail in later chapters. When evaluating the geographic aspects of your study, valuable considerations might include the following:

- Where does your research occur geographically?
- How would you offer a physical description of your study site(s)/research area (i.e., what is the geographic boundary of the   study)?
- Which urban areas or cities are in close proximity to your study site? Are these relevant variables?
- Are there emotionally, spiritually, environmentally, physically, or culturally important units of analysis (neighborhoods, counties, valleys, mountain,)?

- What are the different geographic locations associated with your independent variable(s)?
- Are there any unique natural geographic features that exist in your study site(s)?
- Are there any unique man-made geographic features or facilities that exist within your study site(s)?
- Are there any geographic features of the area that hold some special meaning for the people who live there?
- Are there boundaries associated with your data? How are these defined?
- Are the data stationary through the time frame of your analysis or do they move? How will you account for these changes?

Questions for Analysis

The following list of questions will aid in conducting your analysis. They focus on the issue of boundaries, time and place. All of these issues are discussed in greater detail in later chapters, especially Chapter 6, but are important to have in mind early on. A GIS is useful for conducting a comparative study of groups of people. It's up the researcher to determine boundaries for the study. Boundary choice is essential because changing the boundaries of an analysis can drastically alter the results obtained. Therefore, the boundaries of your study are best dictated by the research question you are investigating and should be set in advance as opposed to during the analysis phase.

Trying to determine whether or not geographic variables will be useful to your study can be tricky, but keep this in mind: Almost all units of analysis have an associated geographic location. Social scientists are often interested in measuring basic sociodemographic variables, such as age, income, gender, ethnicity, and so forth. These core variables can then be associated with geographic locations, such as neighborhoods, cities, and states. The possibilities are limitless. Anything that involves geographic boundaries can be studied using GIS. Sometimes boundaries are artificially created by the researcher, depending on the researcher's goals. At other times, the boundaries already exist and are based on man-made (e.g., streets) or existing (e.g., rivers, lakes, and mountains) geographic features.

The issue of boundaries is a very important one for GIS. One thing to keep in mind is that you are the boss when it comes to your data collection and the type of information that you want to include in your study. Therefore, the notion of appropriate boundaries in your study may require you to look back at your research questions or hypotheses. It is also important to consider whether or not you are dealing with socially constructed boundaries, such as political boundaries (e.g., city limits or state lines) or physical boundaries (e.g., mountain ranges, oceans, or rivers). Sometime the two coincide, but often they do not. Sometimes we use boundaries defined more fluidly and conceptually (e.g., a traditional hunting ground, the heart of the community, or a gang's territory).

Questions for analysis include the following:

1. What is the primary unit of analysis in your study?

2. What are the physical boundaries for your unit of analysis?

3. What are the social (conceptual) boundaries for your unit of analysis?

4. Are your conceptual boundaries social, philosophical, or economic? Conceptual boundaries can be anything that reasonably organizes people into groups relevant to issues you are examining. Although conceptual boundaries are often physically defined, geographic boundaries (e.g., cities or neighborhoods), such boundaries may be self-defined, such as people who identify themselves as part of a particular group because of some perceived shared characteristic (e.g., a particular ethnic or religious group or a group of people with similar a similar hobby or interest, such as gardening or bird watching). You can identify the boundaries of the group you are studying based on these conceptually defined parameters.

5. What is the hypothesis or driving research question? What is the problem or issue your research addresses?

6. What is the main geographic feature or variable you will examine in your study?

7. How is this feature or issue related to your independent or dependent variable? (Also see Question 8 for assistance with answering this question.)

8. Can you identify any geographic pattern by unit of analysis (e.g., group, neighborhood, county) relative to the topic or issue under study? For instance, if you are studying poverty, are there certain neighborhoods in the study that are more in poverty-stricken than others?

9. If your study involves a comparison of groups, are there differences or stratifications based on geographic location?

10. If your study does not indicate any geographically based patterns, are there geographic variables worthy of exploration that you failed to consider?

11. What themes emerge from your data? Do they emerge by unit of analysis?

12. Are there themes that emerge specific to certain geographic locations?

13. Are there clusters of specific social, economic, and political data that appear to be grouped together geographically?

14. If this information is not easily available in a spatial database format, could it be easily created?

How To: Steps in the Process

If you decide that you want to employ a GIS in your project, you need to do two things right away: identify the geographic region and develop a data dictionary.

Identifying the geographic region (or spatial extent) and features important to your study is the first thing you should do. This task includes identifying the categories, geographic features, and physical environmental features that you want to consider as a part of your study. Basically, it means drawing a study site boundary, having a good understanding of your topic, and determining which geographic features you want to measure.

Developing a data dictionary means developing a set of definitions and criteria that for your particular categories. In other words, you need to develop descriptions that clearly define each of the attributes you will be recording in association with your variables. We explore research design and considerations for the incorporation of GIS further in Chapter 4.

For example, you might be doing a study of urban communities, but how exactly will you define these? You might use a definition based on population density, the percentage of the landscape covered by concrete, or the number of roads per square mile. Developing and documenting clear definitions for each variable and attribute to be included in your study are essential steps before going forward with data collection and analysis. Doing this in advance helps to ensure that everyone involved in the study understands the data to be collected and that each item will be mutually exclusive, thus avoiding confusion later in the process.

Public Health Example

The best way to determine the purpose of your research is to think about the central question that you want to have answered. For instance, maybe you are interested in examining whether people who live in a certain section of town (e.g., the poorer section) suffer from respiratory problems. The different parts of the town may be described as being "rich," "middle class," or "poor" based on the income level of residents who live there. The question of health connected to income is a question that could be addressed using a GIS. Following the previously mentioned guidelines for integrating GIS into a study, you would determine the study boundaries and relevant geographic features that are going to be a part of the study. In this case, you are interested in drawing your study boundaries based on income. In other words, you want to draw a boundary around the low-income section of town, the mideconomic section of town, and the upscale part. Where would you begin? You could begin by looking at the U.S. Census data for that town to determine the different clusters in town based on household income. You could then draw your boundaries based on the clustering observed in the

U.S. Census data. Step 1 in the process also calls for identifying other geographic features that might be important to your study. In this example, these geographic features would include the locations of different factories, incinerators, and other production facilities.

Step 2 in integrating GIS into your study calls for developing a data dictionary. You as would determine what level of household income fits your categories: for example rich might be over $45,000 per year, middle class might be $20,000–44,999 per year, and poor might be less than $19,999 per year. In a GIS you could also gather information on the level of emissions from the different facilities by looking at the U.S. Environmental Protection Agency's Web site. On that Web site, the agency indicates industrial sites that emit beyond a certain specified level. This valuable information regarding air quality could become a part of your GIS database.

You could interview people who live throughout the town and conduct a survey that inquires about their general health and income and specifically questions whether or not they suffer from any respiratory problems. As long as you know the geographic location of respondents, you can enter this information, along with the survey answers, into a database. That way, when you conduct your analysis, you'll be able to have geographic information for each unit of analysis (i.e., household). If you want to aggregate the data slightly to protect the privacy of the individuals, you can categorize respondents as living in a neighborhood rather than use actual street addresses as a categories. Upon further investigation, you may want to geographically locate various factories, incinerators, or other production-oriented facilities that could be emitting substances that affect peoples' respiratory health.

Another important thing to consider in determining your research purpose is the general theme that is a part of your research question. In the respiratory health example, some of the potential themes might be environmental health, social inequality, or environmental justice.

After determining the general purpose of your research, you can then ask the question, how would GIS would be helpful to the project? In other words, how would using a GIS enhance the study? A GIS is useful because it facilitates a more holistic and contextual view of the research problem or issue. It accomplishes this by bringing together a variety of different data types. Any study that you choose to develop will most likely include a variety of important variables. The trick in using the GIS is to identify which variables will best be studied using a GIS. This topic is discussed in greater detail in the following chapter.

Relevant Web Sites

"Demystifying the Persistent Ambiguity of GIS as 'Tool' Versus 'Science'": An article by Dawn Wright, Michael F. Goodchild, and James D. Proctor. http://dusk.geo.orst.edu/annals.html

"Social Sciences: Interest in GIS Grows": An article by Michael F. Goodchild of the Center for Spatially Integrated Social Science. http://www.esri.com/news/arcnews/spring04articles/social-sciences.html

U.S. Department of Labor: Bureau of Labor Statistics, *Occupational Outlook Handbook,* Social Scientist section. http://stats.bls.gov/oco/ocos054.htm