Fundamental Research in Geographic Information and Analysis

NCGIA Technical Reports, 1988–1997

University of California,
Santa Barbara

State University of New York
at Buffalo

University of Maine

National Center for Geographic Information and Analysis

Funded by the
National Science Foundation

CD produced with support from
Environmental Systems Research Institute, Inc.

NCGIA

National Center for
Geographic Information and Analysis

Instructor's Guide for

GIS Laboratory Exercises: Introduction to GIS

2nd edition

by

Jeremy Taylor, Jane Fletcher and Karen Kemp
University of California, Santa Barbara

Technical Report 96-12 (supplement)

December 1996

Preface

The following materials have been prepared to assist instructors who are supervising labs based on the materials in NCGIA Technical Report 96-12, GIS *Laboratory Exercises: Introduction to GIS*, 2"d edition. Included are hints on how to prepare for the labs, some guidelines for instructing students through difficult portions of the material and suggested responses for the questions included in the student manual.

Data for labs 2 to 4 can be downloaded from the NCGIA's anonymous ftp site at ftp.ncgia.ucsb.edu in the directory pub/Publications/tech-reports/96/96-12/.

TEACHING ASSISTANT INSTRUCTIONS
Microsoft Windows

The software chosen for these laboratory exercises (ArcView 2.1 and IDRISI for Windows) are both Microsoft Windows based applications. The following is a list of basic tasks the students must be able to perform to complete these labs. It is the T.A.'s responsibility to know these tasks and teach them to the students. **Note:** There is a Windows tutorial under HELP in the Program Manager.

Students must be able to:

- Turn computer on and get to a DOS prompt.
- Start up Windows.
- Use the mouse.
- Get to the Program Manager.
- Run and use the File Manager.
- Navigate drives and directories in the File Manager and within an application.
- Run the software.
- Re-size windows and minimize / maximize applications.
- Use menu, button, and tool bars.
- Save work on floppy disk, in accounts, or both.
- Exit an application.
- Exit Windows.

TEACHING ASSISTANT INSTRUCTIONS & ANSWERS
Lab 2

SET-UP INSTRUCTIONS

Describe the difference between nominal and ordinal data.

TEACHING TIPS

Bring a wall map of Africa to this lab.

ANSWERS

1 . columns = 240: number of columns in the image
rows = 219: number of rows in the image
reference system = lat/long: the status bar will display lat/long coordinates when the cursor is passed over the image
unknown resolution: the cell size (amount of ground area covered by the cell)
12 legend categories: the number of categories/classes into which the data has been divided.

2. object type = line: the vector object is a line reference units = degrees: the units of the coordinates displayed in the status bar.

3. COR or coarse

4. Because the cells are rectangular.
Z:)

5. 20 min. X 20 min. or 1/3 deg. X 1/3 deg. or.333 deg X.333 deg

6. The values are whole numbers corresponding to the legend, they tell you the category number in that cell.

7. The values you get are temperature in Celsius X 10, they mean that the average temperature in January at a cell is...

8. A value of a cell in soiltex refers to a category, or class. A value in tmpjan refers to the actual temperature.

9. Continuous data = elev or others; Classified data = wind or soiltex; Nominal landuse or veg.

10. A palette is used to display different types of data.

11. Least useful should be qualitative 256.

12. Qualitative 256 is more appropriate because nominal data is better displayed with random colors.

13. Answers will vary.

14. The answer should be around 1360 meters. It may be the actual elevation at that point, but it is an average so more than likely it isn't.

15. Because the cells are based on lat/long and the area of a cell based on lat/long decreases as you move toward the poles.

16. The area is 1370 kin. squared.

17. The area is 882 km squared.

18. No, they are not the true area because the area is calculated on a square and the cell is irregularly shaped.

19. Not very accurately, answers will vary.

20. Not very appropriate.

21. North East, mountains, valleys, continental shelf, The Rift Valley .....

22. The vegetation changes in direct correlation to elevation and there are clear latitudinal belts.

23. Most: tree savanna. Least: oasis/rice.

24. Most: tree savanna. Least: marsh/swamp

25. Unit measured is cells. This suggests that when using cells of varying size, area calculations must be carefully done.

26. Class one is most frequent. Class one represents coarse soil texture.

27. Answers will vary.

TEACHING ASSISTANT INSTRUCTIONS & ANSWERS
Lab 3

SET-UP INSTRUCTIONS

none

TEACHING TIPS

Show students how to construct a flow chart
Explain overlay and buffer
Explain the concept of diagonal links (check Idrisi help under GROUPS)

ANSWERS

1.

forspec	shorline	roads
category 1: not coded	category 1: not coded	category 1: roads
category 2: White Pine	category 2: land	
category 3: Jack Pine	category 3: water	

2. 71750 m/205 col = 350 m, 28000 m/80 row = 350 m. Therefore 350 m by 350 m.

3. 71750 in east-west, 28000 north-south. 2009 km^2.

4. Packed binary.

5. Binary images are convenient because they produce output which is easy to interpret by using mathematical operations. Also, the *OVERLAY/MULTIPLY* is equivalent to a Boolean "and" operation.

6. Answers will be something like the final flowchart, but in paragraph format.

7. Multiply, because when multiplying binary zeros and ones, the output is zero which is what is required.

8. **drypine** shows I's where areas are suitable and O's elsewhere.

9. A contiguous stand is the equivalent of a set of cells all having the same value.

10. The cells in each group are given a unique value

11. 38 groups. Use *DISPLAY* or *DESCRIBE.*

12. Assigns each cell a value equal to the area of the group to which it belongs.

13. It covers a large area.

14. It is not a forest stand.

15. The area values, when multiplied with the O's of **drypine,** result in 0 values. The I's result in the same area values. Thus, the background area values are eliminated.

16. You could remove the background areas by doing a reclassification.

17. Assigning, a unique number to each stand allows the stands to be more easily identified.

18. 6 suitable leasing stands.

19. If you include diagonal links, cells will be considered to belong to the same group if they have the same value and they touch in any of the eight possible directions, n, e, s, w, nw, ne, sw, se. If diagonal links are not included, pixels will be considered to be part of the same group only if they have the same value and touch in one of the cardinal directions, n, e, s, w. The results can be remarkably different. In general you want to include diagonal links unless you are trying to uncover polygons separated by lines of a single pixel width. Therefore, if there are cells that would be considered grouped under diagonal links then the number of groups would decrease and the area of the groups would increase.

20. Using diagonal links for leaseable cells may be appropriate. If you want to be sure to include all possible stands for a group (overcompensating) then using diagonal links would be appropriate.

21. 1268 cells fit the criteria. Thus, 1268 * (350 * 350) = **155,330,000 square meters.**
Percentage of White Pine: 5469 White Pine cells only. Thus, 5469 * (350 * 350)
669,952,500 sq. meters. Thus, (155,330,000/669,952,500) * 100 = 23%.

22. **roads**, so that it will overlay the leases since it is a line of cells and would be obscured otherwise.

23. Can't see the roads clearly because they are the same category as the first lease stand. You should change the roads category so that **roads** will have a different value from any values in **lease** and will not be obscured in the final map.

24. Flow chart.

25. Answers will vary depending on the changes in parameters.

26. Major impact from small changes.

27. Very sensitive. That is, any error or the use of poor parameters would cause a huge change in the results. Furthermore, even slight data inaccuracy would also cause a huge change in the results. Thus, it is hard to rely upon the results with great confidence.

TEACHING ASSISTANT INSTRUCTIONS & ANSWERS
Lab 4

SET-UP INSTRUCTIONS

none

TEACHING TIPS

Explain how to create a flowchart

ANSWERS

1. High: central, Ivory Coast; low: Sahara, SW coast, E coast.

2. Nile river delta. Z approx. = 92. Linear feature is the Nile flood plain.

3. Features stand out because they are moist, well-irrigated agriculture regions surrounded by desert.

4. Desert and tropical rainforest don't vary much (they stay consistently high and low, respectively). Grassland and other vegetation types which are responsive to seasonal change (deciduous forest) show lots of seasonal variation in VI values.

5. Bipolar palette shows values either side of zero as ramped colors either side of zero.

6. VI 85>VI 88, therefore with high positive values, VI is decreasing over time. Low values mean vegetation is increasing,

7. There is vegetation in the ocean which is subject to the effects of currents, seasonal change, etc. Also atmospheric effects (clouds) might be causing this.

8. Atmospheric effects, image misregistration, satellite sensor drift, etc.

9. High values on augdiv mean there is less vegetation in 88 than 85. Low values mean there is more vegetation in 88 than 85.

10. Values are divided rather than subtracted, so get a ratio rather than a difference. It conveys change information as does differencing, but the ratio result is easier to interpret as there are no negative values, only amount of change from low values to high values.

11. The magnitude becomes lower as the values go towards zero - it is not just straight difference change. This may be beneficial if the data range is large. For example, vegetation change between 5 and 10 units may not be equivalent to vegetation change between 95 and 100 units. The first is a doubling of the index, the second is a small increase in an already large index.

12. Ratio because it is a ratio (percentage change, though not actually multiplied by 100) rather than an actual values change, so it tells more about the pure change without having to know what the original values were.

13. No definitive criteria have been identified. This should make them think.

14. Threshold value: approx. 35.

15. Not evenly scattered. Most are near coast. Various latitudes. Just south of the Sahara.

16. Forest, grass/shrubland, others.

17. Desert areas don't vary much and thus are unlikely to have a change value above the threshold.

18. Various locations.

19. Flowchart

20. 14 is the number of images in the sequence. 0 is the time delay to use when displaying the images as a sequence.

21. Vertical axis shows the mean VI values for areas of potential deforestation.

22. Over time the VI values have been decreasing, with definite and drastic seasonal variations. The most major change is between August 85 and August 88.

23. August 88 is low.

24. The results were effected because August 85 skewed the output and this cannot be detected until the years in-between 85 and 88 have been seen. When using data with relatively low temporal resolution you must be careful when drawing conclusions about the interceding time.

25. They show similar fluctuations.

TEACHING ASSISTANT INSTRUCTIONS & ANSWERS
for Lab 5

SET-UP INSTRUCTIONS

Create a Proiect & Name it:

Name a view "Conterminous United States"
From data location:_____
Add themes:

- cities
- from USA add region.state (i.e. not region.cnty, nor polygon, nor arc, nor labelpoint)
- from roads add arc
- from waters add region.lakes

(NOTE: click on the folder icons beside the theme names in the add theme dialog box to see the specific themes required)

Customize the Tables

Bring up the tables to all themes.

- Add "stdempct.dbf'and "cidemog.dbf' tables to the project
- Join stdempct.dbf (source table) to Attributes of State (destination table) using common field state-name
- Join cidemog.dbf (source table) to Attributes of Cities (destination table) using common field state-city
- Delete stdempct.dbf and cidemog.dbf from the project

Customize the Themes

Clip Hawaii and Alaska data out of all themes and change cities to only show those cities > 100000 as follows:

- State Theme
  - go to THEME/PROPERTIES/DEFINITION and use query builder:
    ([State-name] <> "Alaska") and ([State-name] <> "Hawaii")
- City Theme
  - ([State_name] <> "Alaska") and ([State_name] <> "Hawaii") and
    ([Type] = "city") and ([Pop1990] > 100000)
- Roads Theme
  - Change name of theme to **interstates**
    ([Admn_class] = "Interstate") and ([Roads-id] <> 169) and ([Roads-id] <> 90

(NOTE: roads 169 and 90 are the ones in Hawaii)

Put the themes in the following order from the top: State, Lakes, Cities and at the bottom, Interstates. Zoom to the extent of state theme, so that the map fills the view. Turn off all themes (i.e. leave check boxes blank). Make all themes inactive (if only one is left active, use shift and click on the theme to make it inactive). Close all tables. Leave the view active and enlarge it to fill the ri-ht hand side of the screen without overlapping the project window.

TEACHING TIPS

You may want to explain -ID and -# numbers in tables in more detail
Emphasize the relationship between the active state of a window and the available functionality
Explain scale ratio representation, the difference between large and small scale.

ANSWERS

1. 29760021

2. 638800

3. Lake Okeechobee

4. The scale becomes larger (i.e. more detailed) on zooming in, vice versa on zooming out

5. The scale ratio represents real map ratio scale. That is, 1:24 000 means one thumbnail on the screen = 24 000 thumbnails on the ground

6. Grouping along Californian coast and eastern seaboard

7. Problem display presents is that city points overlap to make large conglomerations - can't distinguish individual cities

8. Because it is pointless to display cities points on a small scale map since they all overlap
9. Zoom in/out of center of view, spy glass in/out and box select, setting, scale directly using, scale in tool bar, extent of active themes/selected sets in active themes

10. 1:9,999,999

11. Answers may include: data storage reduced (only integers used as in ID field rather

than whole name and address), update only one record and other related records change automatically, data inconsistency problems simplified

12. **Point** table: Area, Perimeter, Cities#, Cities_id. **Line** table: Fnode#, Tnode#, Lpoly#, Rpoly#, length, Roads#, Roads-id. **Polygon** table: Area, Perimeter, State#, State-id.

13. Answers will vary

14. 4/49

15. Answers may vary

16. 16

17. 'New set' selects from a whole set, 'add to set' adds the new selection to a set previously selected

18. 'Select from set' reselects from the currently selected set.

19. States used will probably be from the following set: Oregon, Idaho, Wyoming Utah, New Mexico, Colorado, Kansas and Texas

20. Answers will vary

21. Answers will vary

22. Answer will vary

23. Answers will vary

TEACHING ASSISTANT INSTRUCTIONS & ANSWERS
Lab 6

SET UP INSTRUCTIONS

Before distributing the lab write the path by which to access the data in the space(s) provided.

TEACHING TIPS

Clarify point-in-polygon operations. How they work (i.e. what the computer does to achieve the results). What choices ArcView offers (e.g. "have their center in", "completely contain", "are completely within", "contain the center of" and "intersect").
Clarify joining tables and the relationship between the tract theme and the demographic tables.

SPECIAL INSTRUCTIONS FOR PROBLEM 2

Discuss with the class which logical conditions are appropriate - brainstorm with class etc.
The query works best if the hotel is sited in areas where there are other hotels, as is the case in reality.

ANSWERS

1. Answers will differ

2. From: avdata/atlanta/region
   - shopping centers
   - routes
   - tracts

3. If it is not done the computer will have no concept of distance when doing a spatial query

4. So that there is some reference to tie the two tables together or there is no way for the computer to know which features and corresponding data are the same.

5. Join is "one-to-one" and puts 2 tables together, then turns off the common fields so they don't show twice. An example of Join is if you join a county attribute table to a county demographics table. Each row in the new table corresponds to one record in each table. Link is "many-to-one" and the tables are not physically put together. For example, if you link banks to counties, when you select a county in the counties

attribute table, all the banks in that county in the bank attribute table will become highlighted.

6. A buffer operation in raster selects cells within a chosen distance based on the cell size. In vector a buffer operation a "line is drawn" around the buffered object and features within that area are selected.

7. Answers will vary

8. Answers will vary

9. Answers will vary

10. Water, near other hotels, near a route, standard of living, etc.

11. Answers will vary

12. Answers will vary

13. Answers will vary

14. Answers will differ

15. More realistic: give weighting to parameters

16. Location of golf courses, theme parks, available building sites etc...

17. Limitations: GIS doesn't provide all artificial intelligence - i.e. only as good as the data Capabilities: fast, realistic queries on data given

18. Yes, with very small cell size and many layers, but easier/better with vector

TEACHING ASSISTANT INSTRUCTIONS & ANSWERS
Lab 7

TEACHING TIPS

After editincr a table (in this case adding a record), that edited table must be re-geocoded, subsequently creating a new shapefile. This is because the newly added record in the old table will not appear on the view and the records will be incomplete.
For more information direct students to an Internet URL for Atlanta.
You may need to explain the concept of "diffusion".
Non-topological issues

ANSWERS

1 . You can then carry out address matching.

2. Street network. Others = any network, e.g. sewer lines, town boundaries

3. Process comparing 2 addresses to determine whether they are the same.

4. To tie in the new record to the street network.

5. Answers will vary

6. Answers will vary

7. One way street location, rush hour street blockages, etc.

8. The address matched locations represent a mathematical calculation of the address number, along street segments and does not record offset from the street, which can be significant on large lots. The GPS values, latitude / longitude coordinates, are more accurate because the receiver was placed on the actual location.

9. Data to illustrate diffusion of banks: banks theme, attributes of banks table and classifying according to year established.

10. Use tract demographic table joined to tract theme table. Create field for growth rate 80 - 93 by using ArcView's calculate function.

11. Quantile: equal number of records in each class, Equal Interval: equal range of values in each class, Unique Value: every unique value is put in a separate class.

12. Answers will vary

13. Population growth is positively correlated to bank branch establishment.

14. Similar to all the banks, i.e. well distributed over the region.

15. NW Marietta St. branch because it has the most deposits of the chain and is one of the oldest established ones and is downtown.

16. In the center of the region.

17. Few people live there, negative population growth, few households,

18. Yes. It is a downtown working environment - central business district.

19. Downtown. Tells us that most ATM and branch office deposits are credited to the main branch.

20. Year acquired, but the correlation, although fairly strong, has some anomalies.

21. Economic boom, corporate merger with another bank...

22. Yes, expanded away from the Central Business District as population expanded to the suburbs.

23. Yes, answers will vary.

TEACHING ASSISTANT INSTRUCTIONS
Suggestions for a final exercise

OBJECTIVES

The large landuse and hydrological database for the Bermudez River watershed in Costa Rica provides an opportunity to create an exercise in which students can use both Idrisi and ArcNiew in an integrated manner. Using this database, a final student project can be constructed which utilizes the strengths of each software product and data model individually, and then combines the results of the analyses for a final report.

The database suggests a wide range of possible exercises. Please feel free to download the data and consider how your students might be instructed to use it for a final exercise for your course. We welcome the submission of specific exercises you have produced based on this data. Contact information is shown at the end of this section.

The following section outlines one assignment concept which we hope will inspire the development of your own.

EXAMPLE ASSIGNMENT

The city of Heredia, a provincial capital of a province of the same name in Costa Rica, has just completed a study that cites a 25% deficit in water supply for the projected population growth by the year 2010. To address this problem, an additional water source must be incorporated into the potable public water system of the region and an appropriate treatment plant built to enable the municipality to meet the projected water demands on into the next century.

You have been contracted as one of a set of local consultants to help select a new surface water supply in the Bermudez River Watershed above the city of Heredia in the Central Valley of Costa Rica. There are several potential capture points located in micro-basins of similar surface water discharge. You are charged with examining the feasibility of using one of these potential capture sites.

After determining the types of land use found upslope from "your" capture point, you will identify the potential water contaminants often associated with these landuses. The treatment parameters required for the identified contaminants will determine what type of treatment plant will be necessary to accompany that capture point and will provide the criteria for the selection of an appropriate location. Water will be fed by gravity from the capture point to the treatment plant. Therefore, the facility can be located on any suitable size site which is downslope from the capture point and upslope of the municipal potable water collection and distribution tanks.

DATA FOR THE EXERCISE

The Bermudez River watershed, approximately 40 kilometers squared, is located in the north central section of the Central Valley of Costa Rica. There is a high diversity of land use throughout the drainage basin, ranging from primary virgin cloud forests to light industry such as Firestone Tires, Intel and Craft manufacturing companies, to mention a few. This region is also known for its plentiful ground water supplies that are generated from recharge areas high in the watershed that provide flow for two to three separate regional aquifers throughout most of the basin depending on one's location. The Costa Rican government is banking, on the use of these larger aquifers to supply drinking water for the relatively high population density residing in the western half of the Central Valley where over 65% of the country's population lives.

A digital hydrologic spatial data base has been created to study, preserve and manage the existing hydrologic resources. The data base was created in Arc/Info and the data layers were developed using a 1: 10,000 map series for the most part and occasionally 1: 50,000 map series in the very upper sections of the watershed. Below is a list of the primary data layers available in the data base. Detailed metadata for each data layer is available:

- Drainage network
- Stream flow data
- Isohyets
- Network of meteorological stations
- Recharge areas
- Sub-watersheds
- Aquifer extent
- Equipotential lines for two aquifers
- Contour lines (10-20 meter intervals)
- Land use
- Fresh water capture points (wells, springs and rivers) for potable water as well as agriculture
- Geology
- Soils (note: digitized from a map at scale 1:200,000)

For more information about this dataset and how to download it, contact the following:

Timothy H. Robinson
Project Coordinator
National University of Costa Rica
email: robinson@irazu.una.ac.cr
URL: http:www.una.ac.cr/geog/lsigae

Information about downloading this dataset can also be found in the NCGIA's ftp site at ftp.ncgia.ucsb.edu in the directory pub/Publications/tech-reports/96/96-12/.