Lab 4. GIS Databases and Attribute Tables. Queries.

The goal: To learn about the GIS database management and query..

Objective: Using query language (logical operations) solve two problems..

Software for the lab: ARCVIEW.

Reading: S.Aronoff. GIS: A management Perspective. Chapter 6. Data Management.

---

1. TABLE MANIPULATION WITHIN ARCVIEW

TABLES DEFINITION:

- Numerical or text string information organized in:
  - fields (vertical columns, a.k.a. items).
  - records (horizontal rows <-->).

TABLE SOURCES

- Two general categories of source data for ArcView tables:

1)Theme attribute tables:

- INFO files from ARC/INFO coverages (PAT or AAT).
- Dbase files (.dbf) from ARCVIEW shapefiles.

2)External tables:

- dBase files from other software (EXCEL, Dbase IV)
- Comma or tab delimited ascii/text files (exported from EXCEL or Dbase):
  - First line contains field names.
  - File names must have a “.txt” extension.

VIRTUAL TABLE CONCEPT

- All ArcView tables look and act the same, no matter what the source.
- Tables do not store the actual tabular data, just pointers to the source files.
- Tables can automatically reflect changes made in the source files outside of ArcView.

TABLE PROPERTIES:

- Table title.
- Field position order.
- Visibility of fields. Turn off extraneous fields.
- Alias names to fields.
- Sort values in fields numerically or alphabetically.
- Appearance saved in ArcView project file.

Table Manipulation Exercise:

ADD A THEME AND ITS ATTRIBUTE TABLE

- Open a new ArcView project. Save as "tables.apr".
- Create a new view.
- Add a Basins polygon theme (***/usr/manhattan/gis_class/yxg/data/basin24eoh***) in the view, using the following button:

- Make sure the Basins theme is the active theme.
- Load the theme attribute table:
  - Select this pulldown menu sequence: Theme ---> Table, or press this button:

- Close the view for now.
- Enlarge the table window to see all the fields.

CHANGE TABLE APPEARANCE

- Field order.
  Move [perimeter] field to the far right:
  - Press on the field name, hold, move horizontally.

  Practice on other fields to "get the feel of it."

- Field display width.
  - Move cursor between two field names until <-> appears.
  - Press left mouse button, hold, move horizontally.

- Table Title, Field Aliases, & Field Visibility.
  - Make the table active by pressing in title banner.
  - Select this pulldown menu sequence: Table ---> Properties.
    - Change virtual table Title to: Basin Table.
    - Press into the box to the right of field name.
    - Enter your chosen alias and press enter/return.
    - For this exercise, assign the [Basin] field the alias: Basin_Name.
    - For this exercise, assign the [Area] field the alias: Area_Sq_Meters.
    - Make these fields invisible: perimeter, basin24eoh#, basin24eoh-id :
      - Press on check mark to left of field name. The check mark disappears.
      - If you need to make field visible again, press where check mark was.
  - Press "OK" button when done to dismiss "Table Properties" menu.

- Sort field values.
  - Depress [Basin_Name] field. This is called “making the field active.”
  - Pulldown menu sequences for sorting:
    - Field ---> Sort Ascending
    - Field ---> Sort Descending.
  - Buttons for sorting:

ascending

descending

ADDING EXTERNAL TABLES:

- Comma-Delimited Ascii/Text File
  - Highlight “Tables” on the Project Window
  - Press “Add”. List files of type “Delimited Text”.
  - Select the file “ ***/usr/manhattan/gis_class/yxg/data/data_qual.txt***”.
  - Press OK.
  - This is a table with data collected on Sept 24, 1999, see the following site: http://h2o.usgs.gov/realtime.html.

- dBase
  - Similar procedure to above, but list files of type dBase on “Add” menu.
  - Try adding “***/usr/manhattan/gis_class/yxg/data/basin24eoh.dbf***”.

SAVE THE PROJECT

.

2. QUERYING TABLES

BASIC CONCEPTS BEHIND QUERYING

- Manual or automated search through records in a table.
- Creates a selected set of records.
- Selected set is highlighted in a bright color.
- If theme attribute table, selection and highlighting also apply to geographic features in View.

VISUAL (MANUAL) QUERY

- Pointing to table records.
- If theme attribute table, pointing to theme features in view.

LOGICAL (AUTOMATED) QUERY

- Create logical expression based on field values and logical operators.
- The computer searches the table and creates the selected set based on the logical expression.
- ArcView's logical operators are in Table 1.

Logical Operators

Symbol	Operation
=	Equals
>	Greater-than
>=	Greater-than-or-equal-to
<	Less-than
<=	Less-than-or-equal-to
<>	Not equal
AND	Logical "and"
OR	Logical "or"
NOT	Logical "not"
"*"	Multiple Character Wild-Card
"?"	Single Character Wild-Card

SAMPLE ARCVIEW LOGICAL EXPRESSIONS:

1. ([Basin_Name] = “Kensico”)
Means: select a record with the item [Basin_Name] having value “Kensico”.

2. ([Area] > 5000)
Means: select records with [Area] field value less than 5000

3. ([Basin_Name] = “*on*”) and ([Area] > 5000)
Means: select records with [Basin_Name] containing character “on” and [Area] greater than 5000

4. ([Basin_Name] = “Pepacton”) or ([Basin_Name] = “Cannonsville”)
Means: select records with [Basin_name] containing “Pepacton” or “Cannonsville”

Querying Tables Exercise

VISUAL QUERY

- Have both the basin table and view from the “Table Manipulation Exercise” visible in Arc-View.
- When the table is active, select one or more records with the arrow tool. Hold the [Shift] key down to add to your selection, or remove from your selection. You will see your records in the table highlighted and on the View you will also see highlighted features of your theme.

- To look at all fields for one record in a separate display box use this tool with either a View or table. You will see a separate window with the data for the record.

LOGICAL QUERY

- When the table is active, bring up the Query Builder by pressing this button:

:

- Make sure “Update Values” box on the Query Builder is checked.

Selecting a New Set

- Single click on [Basin] in Fields column and all unique record values for [Basin_Name] will appear in the Values column to the right.
- Double click on [Basin] to put it in the expression box.
- Single click on "=".
- Double click on the value "Kensico".
- Click on the words "New Set" (ArcView highlights the selected set).

Adding to the Selection

- Clear the previous expression. (Press the backspace key on the keyboard)
- Single click on [Sq_meters] (you set this alias for [Area] in the first exercise) in Fields column.
- Double click on [Sq_meters] to place it in the expression box.
- Single click on "<="
- Type the number 60000000 inside the parentheses.
- Click on words "Add to Set".
- You have now added watershed basins with areas less than 60,000,000 sq. meters.
- Move selected records up to top of table. Use this menu sequence: Table ---> Promote. Or press this button:

Selecting from the Selected Set

- Replace the areas criteria (60000000) in the previous expression with 50000000.
- Click on words "Select From Set".
- Repeat process by replacing the area with 15000000. Notice that you get fewer choices.
- Dismiss Query Builder when finished.

ALTERING THE SELECTED SET

- Menu sequence: Edit ---> Switch Selection, or button:

Notice changes.

- Menu sequence: Edit ---> Select None, or button:

Notice changes.

- Menu sequence: Edit ---> Select All, or button:

Notice changes.

ZOOMING TO THE EXTENT OF SELECTED FEATURES
- Use querying techniques to select one of the basins.
- When the view is active press this button and see what happens:

SAVE THE PROJECT

3. EXPORTING TABLES

BASIC CONCEPT:
Tables has to be exported from ARCVIEW when:
1. Table in ARCVIEW has to be used somewhere else as a result of your GIS work;
2. Table in ARCVIEW are “owned” by somebody else and can not be modified until you are the owner; (NOTE: “Own” means ceratain permission level set up within your operating system i.e. read, write, execute. If I own the table you can not modify it. In this case “export table” will basically mean “save table as”).

WHEN TABLE IS ACTIVE WINDOW
- Pulldown menu sequence: File ---> Export.
- Can copy table to another format (INFO, dBase, text).
- Makes editable copies of GIS library attribute tables.
- Makes new table of selected records only. If no records selected than all records are exported.
- Makes new table with visible fields only.
- Makes new table with user-shifted field order.

Exporting Tables Exercise:

EXPORT PART OF BASINS TABLE.
- Open the basin table that you added in the first exercise.
- Unselect all records (a.k.a. “Clear the selected set”).
- Make only these fields visible: [Basin_Name], [Watershed], & [Sq_Meters].
- Shift field order to: [Basin_Name] [Watershed] [Sq_Meters].
- Export the table as a new dBase file in your workspace called “bas_sqmeter.dbf”.
- Add the new table “bas_sqmeter.dbf” to your project and take a look at it.

SAVE THE PROJECT

4. EDITING TABLES

BASIC CONCEPT:

- Editing table means add record or field, change values in records, delete records, fields.
- To EDIT table you have to "own" it, otherwise you will not be able to select editing function. "Own" means that the table should be in the directory that you own i.e. directory where you can read, write and execute. If you do not own the table, then you will be able to access it for viewing only. The "ownership" is a property of your operating system and it is set up according to the operating system rules.

EDITING EXERCISE:

Let's edit the table you created in previous section "Exporting Tables". We'll add another filed called "Sq_Miles" and convert the field "Sq_Meters" to square miles units. The conversion factor between two is 0.000000386 i.e. 1 square meter = 0.000000386 square miles. Do the following:

1. Add the table "bas_sqmeter.dbf" if you do not have it in your project, or open it if you have.
2. Click the menu choice Table ---> Start Editing.
3. Click the menu choice Edit ---> Add Field. Give it the name "Sq_Miles". Leave type as "Number". Make Width = 6 and Decimal Places = 1. Click OK button. You shoud see the field added to your table.
4. Make this field active.
5. Click the menue choice Field ---> Calculate. You will see Field Calculator Menue. Double click on Sq_meters field. Select operation "*" i.e. multiply and type into the window conversion coefficient 0.0000003861. Click OK button. You should see your field "Sq_miles" being populated with numbers.
6. Select Table ---> Stop Editing. Select "Yes" to save edits in your table.
7. Export your new table. Give it the name "bas_sqmiles.dbf". View it.

5. STATISTICS AND SUMMARY

STATISTICS FUNCTION

- Add to your project the table "***/usr/manhattan/gis_class/yxg/data/res24nyc.dbf***". This table describe reservoirs in the Croton Watershed (see http://www.ci.nyc.ny.us/html/dep/html/wsmaps.html). Areas are expressed in square meters.

- Select field "Area" by clicking on it. Select menue choice Field ---> Statistics. You should see the window with all main statistical parameters about your field. Needless to say that "Statistics" will work ONLY with numerical attributes/fields.

SUMMARY FUNCTION

- As you can see in this table, the Field "Name" has many attributes with the same value, i.e. several records contain the same name of the reservoir. This happened because some reservoirs were digitized together with small lakes draining into the reservoir. Therefore, attribute with the reservoir name can be seen few times. In this case if we would like to find the total area of the reservoir we would have to select all records with the reservoir name and then use "Statistics" on selected field "Area". This is a time consuming and not effective way. To deal with this and similar tasks when several records have to be grouped and analayzed, there is a tool called "Summarize" under the Field menue next to "Statistics" tool.

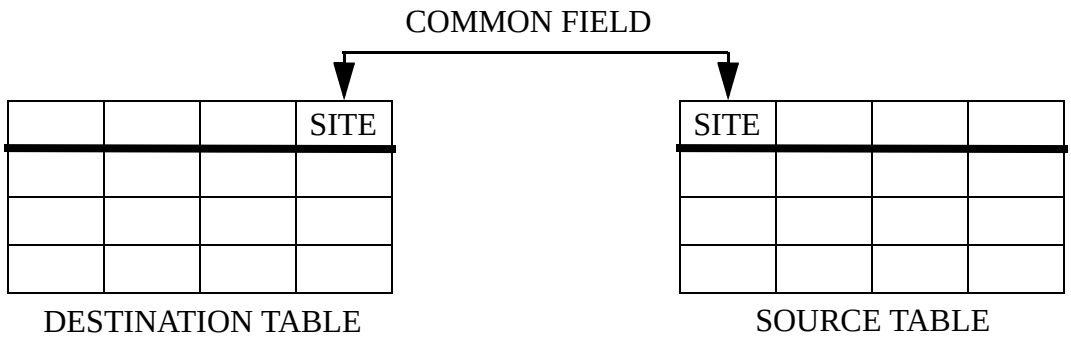
- Make the field with reservoir names active. Select the menue choice: Field ---> Summarize.
- You should see the menue called Summary Table Definition.
- "Save As" res_sum.dbf.
- Select Field called Area.
- Summarize by Sum (you have to select this).
- Click button Add (it will create a field Sum_Area in your saved table res_sum.dbf.
- Hit button OK.

You should see the new table res_sum.dbf with three fields: Name, Count, Sum_Area. As you notice, in the field Name reservoir names are unique. Count shows the number of records from the table res24nyc.dbf which were "summarized" regarding area size.

5. JOINING

BASIC CONCEPT OF THE RELATIONAL DATABASE:

- Table records relate to each other through the values in a common field.



Advanatges:

- Allows for non-redundant data storage.
- Allows for decentralized database maintenance. Different departments/groups are responsible for different tables.

Disadvantage:

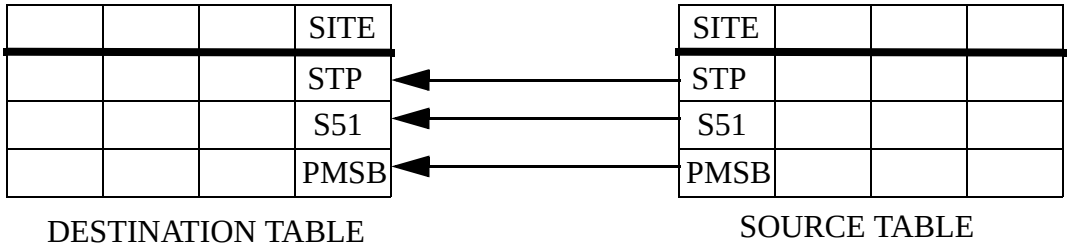
- Does not forgive inaccurate data entry i.e. mismatched characters and numbers.

ARCVIEW TERMS

- Destination Table: Table receiving results of relate. Usually theme attribute table.
- Source Table: Table containing info to be related. Usually an “external” table.

TYPES OF TABLE RELATIONSHIPS

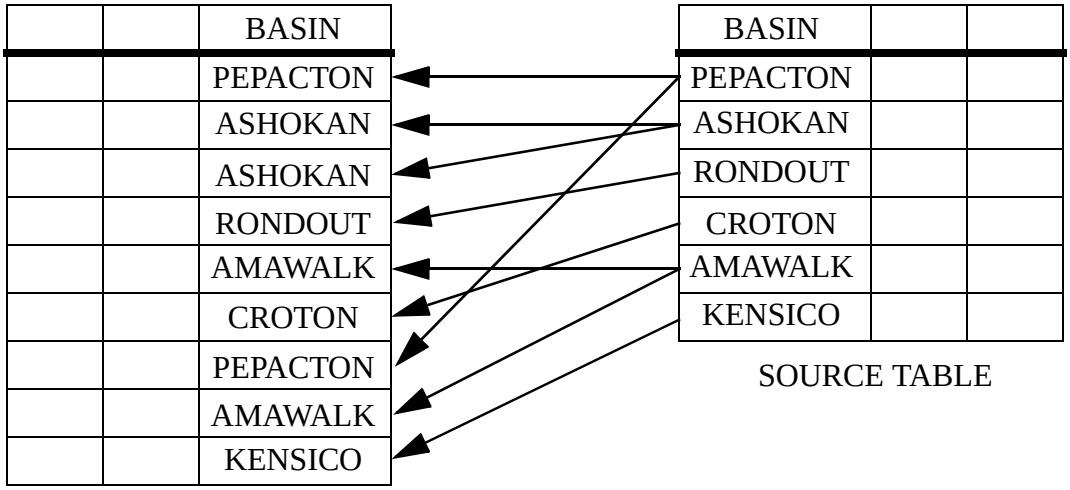
1. One-To-One
   - Each record in destination table has only one match in the source table.



DESTINATION TABLE SOURCE TABLE

2. Many-To-One

   - Many records in the destination table are related to one record in the source table.

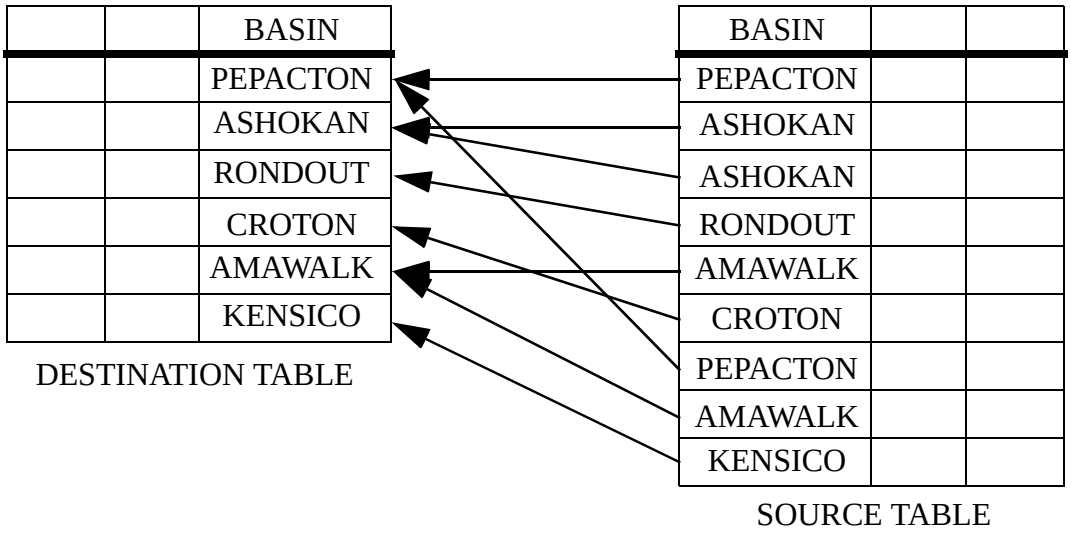


DESTINATION TABLE

SOURCE TABLE

3. One-To-Many

- One record in the destination table is related to many records in the source table.



RELATIONAL DATABASE MANAGEMENT IN ARCVIEW

- **Join**
  - Compatible with One-to-One and Many-To-One, but NOT One-To-Many.
  - **Tables appear physically joined together.**
- **Link**
  - Only way in ArcView to do One-To-Many.
  - **Tables do not appear physically joined together**. (We'll not do exercise on this because it is done the same way as "Join" operation.

ARCVIEW RULES FOR COMMON ITEMS (JOIN & LINK FUNCTIONS).

- Common field name: this DOES NOT have to be the same in each table.
- Common field type: numeric-to-numeric OR string-to-string only.
- Joins/Links are not permanent, but the definition is saved in project file when you save and close your project. To permanently save joins use EXPORT function.
- Multiple joins/links are possible on one destination table.

Joining & Linking Tables Exercise:

ARCVIEW JOIN FUNCTION, ONE-TO-ONE RELATIONSHIP

- Make a new project called "join.apr". Create a theme attribute table for all of the Croton watershed basins (***/usr/manhattan/gis_class/yxg/data/basin24eoh***). This table appears with the default name "Attributes of Basin24eoh". Add an external table with the road mileage for each basin (***/usr/manhattan/gis_class/yxg/data/roads_per_basin.dbf***).
- Close any open views and arrange/resize the tables so they are both visible in the project window.
- Join "***roads_per_basin.dbf***" to "Attributes of Basin24eoh". The common field is [Basin_Name].
  - Destination Table = "Attributes of Basin24eoh"
  - Source Table = "***roads_per_basin.dbf***"

- To join two tables:
  1. Make the source table active.
  2. Click on the name of the common field you wish to use.
  3. Make the destination table active.
  4. Click on the name of the common field you wish to use.
  5. From the Table menu choose Join option, or click the Join button:

E-mail me:

Using sort buttons answer the following question:
Which basin has the most road mileage and which one has the least ?

Use the menu option: Field ---> Statistics. What is the total road mileage for the watershed ?

- Unjoin the two tables by pressing Table menue option "Remove All Joins" when the destination table, "Attributes of Basin24eoh", is active. Close "Attributes of Basin24eoh".

ARCVIEW JOIN FUNCTION, MANY-TO-ONE RELATIONSHIP

Many-To-One relationship in this exercise is used to create a table that would display the area of the reservoir for each record containing sampling site name. Original table with sampling sites does not have the area of the reservoir. Therefore this table will a destination table. Table from the

"STATISTICS AND SUMMARY" exercise "res_sum.dbf" will be used as a source table.

- Make sure you have the table "res_sum.dbf" and "***/usr/manhattan/gis_class/yxg/data/site-lim.dbf*** in your project.
- Join "res_sum.dbf" to "sitelim.dbf" using the fields with basin names as common fields.
  - Destination Table = "sitelim.dbf"
  - Source Table = "res_sum.dbf"
- Query for the reservoir name of Croton. Promote selected records. Note how the [Sum_Area] value from "res_sum.dbf" table is repeated for each record.
- Unjoin the two tables by pressing [Menu: Table ---> Remove All Joins] when the destination table, "sitelim.dbf", is active.

SAVE THE PROJECT.

Spatial Join

- Equivalent to overlaying theme features, but result appears in attribute table only.
- Procedure is same as a "regular" tabular join, except that the [shape] field is the common field in both tables.
- Again, multiple joins are possible.

SPATIAL JOIN TYPES

DESTINATION TABLE		SOURCE TABLE
ARE COMPLETELY WITHIN		
Point	in	Polygon
Line	in	Polygon
Polygon	in	Polygon
INTERSECT		
Line	on	Line
NEAREST*		
Point	to	Point
Point	to	Line
Line	to	Point
* [Distance] field is temporarily added to destination table.		

Spatial Join Exercise

SUMMARIZE New York State Department Environment of Conservation (NYSDEC) WET-LAND AREA PER BASIN.

NYSDEC wetlands are regulated wetlands that has to be protected according to the federal law. The criteria for the protection of wetlands is their size. It should be greater than 12.4 acres.

- Add a labelpoint theme (***/usr/manhattan/gis_class/yxg/data/decwtpl24eoh***) to your view. Using theme properties, call it "DEC Wetland Labels". Each labelpoint represents a center-point of a wetland polygon. We are using labelpoints for this analysis because the spatial join works faster on labelpoints and the attribute table source data is the same as for the polygons.
- Add a theme attribute table called "Attributes of EOH Basins" using basin dataset ***/usr/manhattan/gis_class/yxg/data/basin24eoh***. Assign the alias "Basin_Area" to the [Area] field.
- Get the theme attribute table for "DEC Wetland Labels". In this table assign the alias "Wetl_Area" to the [Area] field. Notice that this table does not have a previously assigned field for basin name.
- Perform a "Point-in-Polygon" spatial join with the following:
  - Destination Table = "Attributes of DEC Wetland Labels"
  - Source Table = "Attributes of EOH Basins"
- To do this, use fields "Shape" from both tables as a common field.
- Examine the resulting table. Notice how each wetland label record has the record of its corresponding basin appended onto it. Make invisible the items you do not want to see.
- Summarize the table on the basin name field. Sum the [Wetl_Area] field as your statistical operation. Name the table "wetl_sum.dbf."
- Add a numerical field called "wetl_acres" (width and decimals your choice) to the table and populate it with the number of wetland acres per basin. There are 4047 square meters per acre. Do not forget to "Stop Editing."

SAVE THE PROJECT.

E-mail to me:

How many NYSDEC wetland acres are in Kensico Basin?

HOMEWORK:

You will have to e-mail me TWO answers from the Lab 4 exercises. Besides this you will have to apply GIS to the following problem:

Problem:

Management of the watershed requires the evaluation of basins where runoff can be extremely high. Special engineering designs have to be applied for these areas to prevent pollution from entering water courses and reservoirs during rain storms. The first step is to define basins where the area of the basin considerably exceeds the area of the reservoir. To evaluate this condition, “reservoir area/basin area” ratio can be used. This universal dimensionless parameter will indicate relative proportion between the area of the dry land and water body.

What to do:

- Using two GIS datasets - basin boundaries (basin24eoh) and reservoirs (res24nyc) define the ratio between reservoir and basin areas i.e. “reservoir area/basin area” . Then sort your data-base and name the first three basins where this ratio is the greatest.
- Describe the sequence of steps you used to resolve the above problem and send me everything via e-mail.