TABLE TIME

(A reading and an exercise)

Background

Two key technical capacities separate those who can succeed in using **geographic information systems (GIS)** and those who cannot:

1. ability to navigate and work with computer drives, folders, and files
2. ability to use data in tables

The first item reflects general computer experience; GIS users must have basic computer skills, and comfort with files and folders is a particularly useful measure. The second item involves the fuel for a GIS; successful users of GIS understand how data can be organized and what kinds of operations can be performed on collections of data within tables.

Thinking about Tables

Let's look at the Smith family.

a) Dad - Patrick Smith, male, 42, baker, 5'10", 170 lbs, brown eyes, likes to fish, graduated from high school
b) Mom - Paula Johnson Smith, female, 43, teacher, 5'7", brown eyes, likes to swim, graduated from college
c) child1 - Tony Johnson (child of Paula's first marriage), female, 16, student, 5'8", 140 lbs, blue eyes, swims, 10th grade
d) child2 - Terry Smith (child of Patrick's first marriage), male, 11, student, 4'11", 100 lbs, hazel eyes, likes computers
e) child3 - Tina Smith (child of Paula and Patrick), female, 7, student, 4'4", 87 lbs, brown eyes, likes to fish

We can think of each person in terms of a description but it's hard to answer some questions, such as

- What is the Smith family's average weight?
- How many eye colors are represented in the family?
- If the family goes to a cabin on a lake, which people are likely to be happy?
- Is there a relationship between physical features and personality traits?

These and many other questions will be much more easy to address if the data about the family members are displayed in a table.

Tables allow users to see information in a "linear fashion" -- similar things organized in lines. You can create a line for each family member, and for each characteristic. But the lines need to be organized in a way to help users understand the information available.

For the Smith's let's put their information into a table, **grouping information by "like things":**

First Name	Last Name	Family Role	Societal Role	Sex	Age	Ht	Wt	Eye Color	Educ	Likes
Patrick	Smith	Dad	baker	M	42	5'10"	170	brown	high school diploma	fishing
Paula	Johnson Smith	Mom	teacher	F	43	5'7"		brown	college degree	swimming
Tony	Johnson	Child1	Student	F	16	5'8"	140	blue	10th grade	swimming
Terry	Smith	Child2	Student	M	11	4'11"	100	hazel		computers
Tina	Smith	Child3	Student	F	7	4'2"	72	brown		fishing

You can look across a row and see a single person in the family. You can look down a column and see a particular characteristic across the family.

In "table talk", the rows are known as **"records,"** and the columns are known as **"fields."** Each individual box is called a **"cell."** Each family member has spaces for the same kind of information. Each record has the same columns, presented in a similar fashion; each field covers an equal number of records.

The fields differ in form, from one to the next, but each column is alike all the way down. Because the columns are alike, we can sort and query, trying to answer some interesting questions, such as:

- How many have brown eyes?
- How many have brown eyes *and* like to swim?
- How many *either* have brown eyes *or* like to swim?
- What are common characteristics about those who like to swim?

However, there is a problem in the data. Some cells are not filled in -- there are no data about certain characteristics for some people. If we want to ask questions about the family, does it matter if there are empty cells? Can we just make up something to put in the empty cells?

Look at the "Wt" field. What is the average weight in the family? Paula's weight is not listed. Should we put "0" in the cell? Would it work to enter "Yes"? For a computerized table on which we want to perform mathematical procedures, "number fields" work best when (and perhaps only when) every cell in the field contains a true number. If each cell needs a valid number, empty cells should be given a real number that, by its nature, indicates something unusual. In the world of GIS, a common strategy is to fill empty cells in a number field with "-99", or some equally unusual number to attract attention and so cause the analyst to exclude the entry from a particular exploration.

Next, look at the "Ht" field. The cells contain numerals plus special characters that are indicators of feet and inches. Tables are great for sorting, but sorting cells that contain any "alpha text" (called a "string field") is done alphabetically. Since these cells have "numerals" rather than "true numbers", sorting alphabetically would list heights (from lowest to highest) 4'11", 4"2", 5'10", 5'7", 5'8". (Notice that this column lines up on the left, in contrast with a number column like "Wt" which lines up on the right.) To change heights into true numbers on which we could sort and query mathematically, we should convert each person's height into total inches (70, 67, 68, 59, 50 from top to bottom of the table above).

Last, look at the field names. In a printed table, it's OK to have column headers like "Societal Role" because it makes sense to us. But some database software will only accept field names that are a maximum of ten (10) characters. Some software requires the field names to begin with a letter, be made of only numbers and letters, with no spaces, periods, or other characters permitted except for an underscore ("_"). In this case, "Societal Role" might be changed to "SocialRole", and "Ht" might be changed to "Ht_inches".

Why tables?

Why go to all this trouble to build tables? Imagine thinking about all the students in your class, all the towns in your region, or all the people in your country. Think about all the characteristics that might describe them. Computers are great tools for sorting and sifting volumes of information about individual items. To do this properly, though, computers need the data in prescribed formats. You can ask very important questions, but the only way a computer will be able to provide a proper answer is if the data are organized and consistent.

By creating computerized geometric shapes that represent features on the earth (countries, states, cities, rivers, etc) and building tables that describe the characteristics of those features in common ways, we can use the power of the computer to analyze information contained in those tables. This is GIS!

TASK #1:
Build a table about the people you live with, or about a family you might imagine. In addition to the name, be sure to include at least one field that requires a single character (like "M/F"), at least two additional string fields, and at least three fields that require a true number entry (like age, height, shoe size, or best bowling score).

NAME	TEXT1	TEXT2	TEXT3	NUMBER1	NUMBER2	NUMBER3

After completing your table, write down some questions that could be addressed by sorting and sifting in the table you have created.

TASK #2:
Build a table of 5 locations around the school, using the following table structure as an example. You can choose different field content for the last two, but make sure you have a field for site (could be name or number), latitude (degrees decimal), longitude (degrees decimal), an additional string field, and an additional number field. (For help on longitude and latitude, see your GPS manual or http://esriurl.com/latlongfinder).

SITE	LAT_Y	LONG_X	SURFACE	TEMP
A1	45.0	-93.3	Grass	20
A2	44.987	-93.299	Pavement	23

TASK #3:
Having built the table in Task #2, create a computer version of the same table. Create it using **TextEdit** (Macintosh) or **Notepad** (Windows), with commas to indicate a jump from column to column. (Do not use a comma anywhere else.) For the example table in Task #2, we would create a file called **"task2data.csv"** that would look like this:

```
Site,Lat_Y,Long_X,Surface,Temp
A1,45.0,-93.3,Grass,20
A2,44.987,-93.299,Pavement,23
```

Notice that there is no space after any comma, and each record is on its own line. When this is all done, you should be able to move your "task2data.csv" into your GIS software (such as **ArcGIS Desktop** or **ArcGIS Online**) using standard procedures, see your data on a map, and analyze it!