

## LAB EXERCISE B: Elementary Database Management

In this exercise, you will create a database from a text file, then run queries on that database to answer questions about Superfund sites in New England. In addition to familiarizing you with the mechanics of Microsoft Access, this lab highlights the importance of being able to bring a raw text file from various sources into a database package (and not just work with data that has already been pre-loaded). Often we can find lists of things on the Internet or elsewhere, but then we need to find ways to bring them into familiar programs for further analysis. Seeing what it takes in this case should help you recognize the issues in parsing the data properly (and defining it unambiguously from the start). We use some EPA (**Environmental Protection Agency**) data on the net to illustrate such a situation in this exercise.

[Images and output created with MSAccess® software.]

1. *Examine the text file:* Use the WordPad program to look at the ASCII text file **C:\11.208\data\sites.txt** and determine how it is formatted. (You can locate WordPad in the Start menu under **Start > Programs > Accessories > WordPad**.) Are the entries delimited by tabs, commas, spaces, or something else? You will need this information later, in step 3.
2. *Launch Access and create an empty database.* Start Access and tell it to create a new database using a blank database. Save it in your personal 11.208 folder and call it **lab\_b.mdb**.
3. *Import a text file into the database.*
  - a) Use the menu item **File > Get External Data > Import** to display the 'Import' dialog box, which lets you select the file to import into the database. Be sure to set the 'File Type' to 'Text Files (\*.txt;\*.csv;\*.tab;\*.asc)' to ensure that 'sites.txt' will appear when you navigate to **C:\11.208\data** folder. Click on the **Import** button, and the 'Import Text Wizard' will appear. This wizard will walk you through the steps of importing the file into a new Access table.
  - b) On the first screen, it will ask you to identify your file as delimited or fixed width. These data are delimited, so choose that (the default), and click the **Next >** button to move on.
  - c) On the second screen, Access asks you to pick the delimiter you want to use: tab, semicolon, comma, space, or something else. Access will make a guess as to which delimiter it thinks is being used. Experiment by picking different delimiters and watch how the preview changes. Since the first row of our file does *not* contain column names, do *not* check the 'First Row Contains Field Names' checkbox. The 'Text Qualifier' can be used to embed the delimiter character within a field. For example, if the delimiter were a semicolon (;) and the text delimiter were a double quotation mark ("), the text

**hi;there**

would be interpreted as two separate fields, 'hi' and 'there'. However, the text

**"hi;there"**

would be interpreted as a *single* field, 'hi;there', even though it contains a delimiter. The file we are working with, however, does *not* use a text qualifier, so leave this set to '{none}', the default.

When you are satisfied with your choice of delimiter, click on the **Next >** button to move on.

d) On the third screen, Access will ask you whether you want to store the data in a new table or add the data to an existing table. Since we have no tables in the database yet, select 'In a New Table' (the default) and click on the **Next >** button to move on.

e) On the fourth screen, Access give you a chance to name your fields and describe their data type (text, integer, date/time, etc.). We will treat all the fields as 'Text'. However, we would like to give them more meaningful names than 'Field1', 'Field2', etc. By clicking on the field in the preview portion of the window -- and highlighting the column -- you can gain access to the Field Options for that field. Rename the fields as follows:

Old Name	New Name
Field1	Site
Field2	State
Field3	City
Field4	Fed_Owned
Field5	State_List

Note that if you don't have five fields at this point, you may have picked the wrong delimiter in an earlier step. You can use the **< Back** button to go back to a previous screen and correct any errors.

Do *not* check the 'Do not import field (Skip)' box for any field, and leave the 'Indexed' property set to 'No' for all the fields.

When you are ready, click on the **Next >** button to move on.

f) On the fifth screen, Access recommends that you specify a primary key for the table. We will discuss primary keys later in the week, but for now the description that Access provides in the screen is sufficient:

*A primary key is used to uniquely identify each record in your table. It allows you to retrieve data more quickly.*

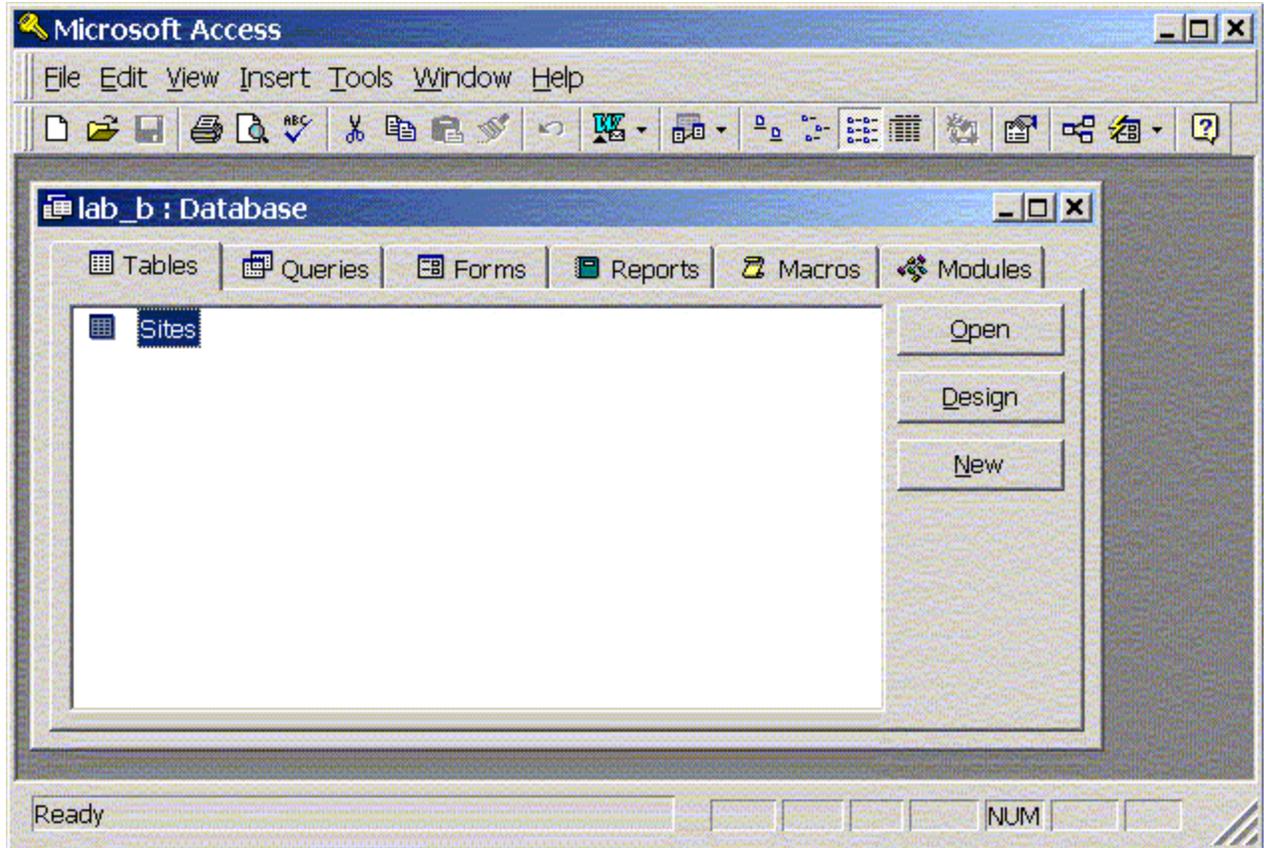
We agree that defining a primary key is a good idea. Access lets you choose an existing column in your table, or you can let Access add one called 'ID' that contains sequential integers. The numeric ID is a good choice, so choose the option 'Let Access add Primary Key' (the default) and click the **Next >** button to continue.

g) On the final screen, Access asks you for the name of your new table. The default, 'Sites', is fine. If you are comfortable with all your choices so far, click on the **Finish** button to create your table. Otherwise, you can use the **< Back** and **Next >** buttons to navigate through the previous screens until you are ready. After you press the **Finish** button, Access should report:

Finished importing file 'C:\11.208\data\sites.txt' to table 'Sites'.

If you see this, it worked!

4. *View the 'Sites' table.* At this point, you should be able to see your 'Sites' table in your 'lab\_b' database. The Access window should resemble this image:



### **Access database with 'Sites' table**

Double-click on the 'Sites' table shown in the window to view the table's contents. A new window called 'Sites : Table' should appear, and you should see all the data in the text file, now nicely organized into rows and columns.

5. *Adding records to the 'sites' table:* Let's add a few new rows of data. Scroll to the bottom of the table, where you'll see an empty row, except for the text '(AutoNumber)' in the first column and an asterisk (\*) in the far left margin. First determine how to encode the following two new sites, and then add the appropriate information into new rows at the bottom of the table. All you have to do is start typing into the empty row. Notice that Access automatically adds the next value of the 'ID' field for you! In order to make query results consistent, try to enter these new records using the same abbreviations and upper/lower case as is used in the rest of the file.

Precision Plating Corp., located in Vernon, CT. This site is not a federal facility, and is not on the state priority list.

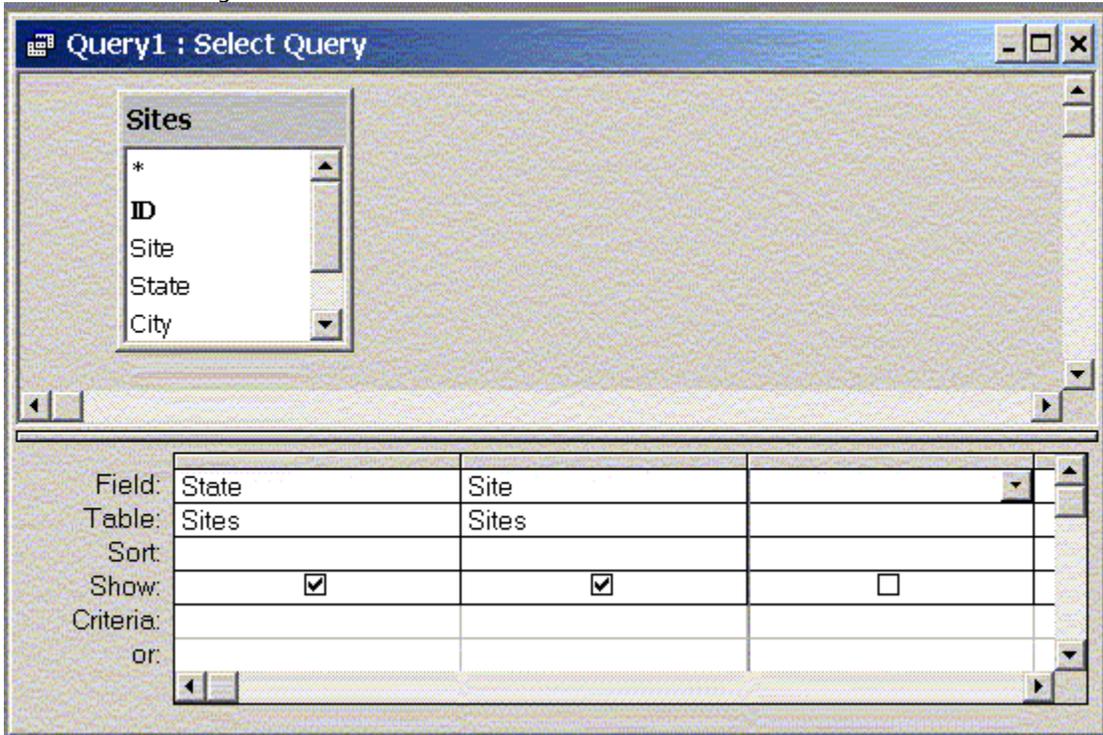
Pease Air Force Base, located in Portsmouth, NH. This site is a federal facility. It is not on the state priority list.

After you've entered these new rows, use the **File > Save** menu item to save your work.

6. *Designing & Running Queries:* To query your table, click on the **Queries** tab and then **New**. At the **New Query** window, you have a choice of Design View or several Query Wizards. For now, choose Design View and click **OK**.

At the **Show Table** window, under the **Tables** tab you should see **sites**. With this highlighted, click on **Add**. You should see a **Sites** window pop up in the **Query 1: Select Query** window. At this point, since you are only using one table for Lab B you may close the **Show Table** window. Note that in Lab C you will need to use several tables and so you will add more than one table.

To conduct a query, you must now select the specific fields you wish to query. You may drag and field name from the **Sites** window and drop it into the Fields List boxes. Alternatively, you may click on an empty Field box and select from the pull down list. For example, set up a query to answer question (a) in Lab Assignment B by selecting fields State and Site and adding them to the Fields List boxes. Your fields list should look like this:



**Select Query Window**

Note that you must now add criteria and sorting to actually answer the question! Refer to the examples shown in the Lecture Notes on the class web site. Refer to Lecture 2: Elementary Database Management.



Remember that you need to click on the Query Now button  to run your query. Also, you can save a specific query and its associated results for future reference. Now answer the questions on the Lab B Assignment Sheet and turn in your answer sheet to a lab instructor or to the 11.208 assignment box in the hallway to the CRL/PSS offices.