Microsoft Access Hands-on CAR Training

Quick Tips

To browse your file, make sure the TABLE tab is selected, then double-click on the file you want to read.

To create a new query (remember, you can only ask questions in a query). Click the new object button and choose NEW QUERY. From the window that pops up choose QUERY DESIGN ('cause query wizards are for wimps).

To run your query, click this button:

To reuse your query -- go back to ask something else click this button:

Every query consists of up to four steps:

1. Select the fields you want to show. You do this by double-clicking on the item from the table list in the top left corner of the query window.
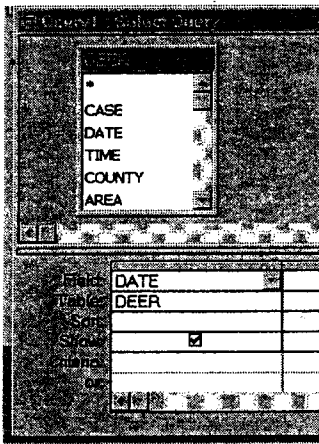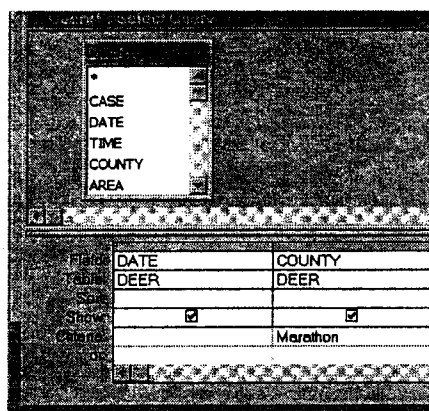
Access will put the field name on the field line of the query grid -- it also checks the SHOW box so you'll see it in the answer.

2. Sort the data. On the second line of the query grid is the sort box. If you click on it, it gives you the sort options.
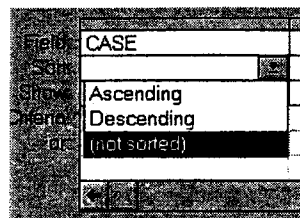
3. Filter the data. If you only want part of your database, you do this using CRITERIA.

4. Summarize your data. If you don't just want to see individual records, click this button:

Access will add a TOTAL line to your query and stick in the nerdy term "GROUP BY." Group by just means you're making piles of certain types of records. Below we've made piles of counties.
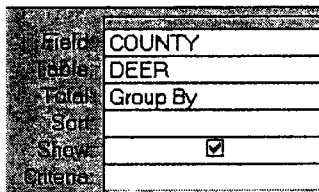
Once you have your piles, you can go back at them and count, sum, average. Here's what you'd do if you wanted to count the number of counties in this database:

Jennifer LaFleur, St. Louis Post-Dispatch
Drew Sullivan

	COUNTY	COUNTY
Field	DEER	DEER
Table	Group By	Count
Total		Descending
Sort	☑	☑
Show		
Criteria		

WARNING: This only counts non-blanks --- if you want to include blanks in your list use COUNT(*) and change your total line to EXPRESSION.

Here's what you'd do to average the ages by county:

	COUNTY	SAGE
Field	DEER	DEER
Table	Group By	Avg
Total		Descending
Sort	☑	☑
Show		
Criteria		

BIZARRE RULE #57: When you want to filter data that is in a summary query, use WHERE on the TOTAL line:

	COUNTY	SAGE	COUNTY
Field	DEER	DEER	DEER
Table	Group By	Avg	Where
Total		Descending	
Sort	☑	☑	☐
Show			Marathon

Joining tables: When you need to join two or more tables, they must have some field in common (it doesn't have to have the same name, just the same data) Click and hold on one of the field names and drop it on the other -- Access gives you this great line to illustrate the join:

Putting the results in a table: Most queries are called Select Queries -- that means you just see the answer on the screen, but it's not saved anywhere. To take the result of a query and put it in a table, BEFORE YOU RUN THE QUERY, click QUERY, Make Table -- you won't see the answer. Access will tell you that it copied so many records to the table.

Importing text files: To import a text file, click FILE/IMPORT -- access will prompt you for the type of file you want to import. Most text files are either FIXED or DELIMITED. Access will then ask you for the file name.

If it's a fixed file, you also will have to create a SPEC file -- basically the record layout for the database. If you don't have that information, you can't use the data. Once, you've typed in the spec, save the spec file -- then Access will do the import.

For more information, get a copy of Importing with Access.

Negotiating for Electronic Information

*You always can get what you want, when you want it
in whatever format you need*
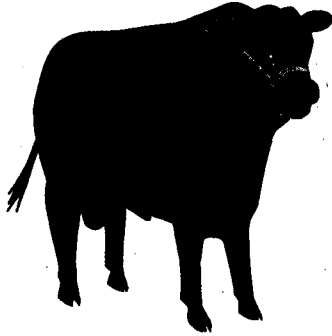
(that's a little bull)

Just like dealing with paper documents, getting electronic information from a government easy can be difficult at times, but by following a few strategies and learning the ins and outs of data, it will get easier.

When you try to negotiating for an electronic database here are some strategies to follow:

1. Know the law. Know how your state treats (or doesn't) electronic information and what the exemptions are.

2. Know what information you want. Don't ask an agency to provide you everything they have. Make sure your request is narrow and specific. Also, you don't always have to ask for the data first -- request a lists of their databases. Then ask for record layouts for specific databases. Then the data. Or from a printed report, find out which databases they use to generate that report.

3. Know how the information is kept. Try to find someone in the information systems department at an agency who knows how they keep the information.

4. Know what the appropriate cost should be. You really should only have to pay duplication costs: have them itemize those costs.

5. Know who does the data entry. The best resource to any database is the data entry clerk. They can tell you things like "oh, we usually just skip that information."

    Know who administers the data. The person in charge of the database can be much more helpful that the PR person. Sometimes they can be pretty excited to have someone take an interest in their database. Long before you ever need the data, go a tour of the agency's data processing center -- get to know the folks you need to meet.

6. Get hardcopy summary reports. This will give you a way to check your data.

7. Know how many records or pieces of information are in the database. When you get the database make sure you have the right number of records.

8. Know how large the database is. You don't have to know a lot of technical jargon. Your PC will hold at least 250 megabytes of information or roughly 250 million characters. If they say the file is 900 megabytes, you won't be able to do much with it unless you find a computer with a bigger hard drive.

9. Don't settle for less. A government agency may claim that certain pieces of information are confidential. Don't just accept their answer: once you do you set a precedent. Go to your editor, call the company lawyers, call IRE's access committee and find out if that information really is confidential.

10. Go to local software users group meetings -- there's usually some data processing folks from government agenies there. (Especially things like ArvView)

Probably two of the biggest issues we face today are privacy and public records for profit. It is increasingly difficult in many states to get any data on individual people -- and there's little public support for access to such data. Be persistent at getting the data while it's still available and get together with other news organizations to fight closing these databases.

When it comes to government agencies making money from data or hiring outside companies to process their data, you again need to be persistent about only paying reproduction costs. Larger news organizations may just go ahead and pay the price to get the data quicker -- but it sets a precident and hurts all of us.

Overcoming Common Excuses

All the previous guidelines are all dandy, until you actually go to ask for the data. Many times you'll run into excuses. But keep in mind 9 out of 10 reporters will just go "oh, okay" and go away when confronted with these excuses. You must not do that! You must keep fighting for it! (Okay, so I'm feel strongly about this issue)

☎ **Listen Missy (I REALLLY hate that one) our database is very complicated you probably won't understand it.**

Oh really, are you using some kind of variable-length or redefined records in your relational database design that won't just export to an ASCII comma-delimited file? If you're not dealing with someone who knows the data, this usually gets you there.

OR... **Our database is very large**

How large?

Thousands of records

Hmmm...my PCs hard drive is 5 gig, you that will hold it?

4

☎ **Our computer system can't do that.**

Really, that's unusual for today's database systems. What system do you use? (This should eventually get you to someone who knows what's going on)

You may also run into cases where they truly do not understand how to do anything with their computer except run the reports they know how to run -- find out who the vendor is and talk with them. I've also run into cases where they used software I was familiar with and could talk them through the process over the phone.

☎ **The person that knows how to do that is on vacation for two weeks /doesn't work here anymore.**

This is a problem. Know one else knows how to do it? What if there were an emergency? This sounds like an even better story.

☎ **It will cost you $20,000**

Ask for an itemized estimate of charges. Three out of four times, this gets the cost down at least some. "Oh we recalculated the costs."

> Offer to pay reasonable programming fees. (see handout on costs) You should only have to pay the programmer, their hourly rate, NOT OVERHEAD.

> IF THEY DO PROGRAMMING -- ask for a copy of the program or at least have them put in writing that they will save the program in case you need the data again next year.

> See if there is a rate charged by state agencies to other agencies.
> Provide your own tape
> Ask for a backup tape.

In some cases this actually might be the price charged to that agency from a larger state data processing center. For example, in California, many agencies are charged such fees from the Teale Data Processing Center. That's an even better story.

A great hint from Dave Armstrong at the Boston Globe: Do a records request for the data they've provided and what they have charged other organizations.

☎ **The database is not public record.**

The burden is on them, not you to show you where in the open records law that information is not a public record. In California, for example, they don't have to give you the data -- so they actually may have an argument.

5

☎ **We don't like what you plan to do with it.**

Interesting. What do you think of some of the other stories we've been doing in the paper? So how soon can I get these data?

☎ **The database contains confidential information**

They must tell you what pieces are confidential. In some states they must redact the information and give you the rest -- the electronic equivalent of a black marker.

☎ **We don't keep that on computer.**

Make sure that's true: Interesting, so how do you keep track of all these people? You do that all by hand?

You may actually run into cases where it's not computerized -- in that case, you may want to have a data entry how input the information. I'd advise against scanning, especially if there are lots of numbers involved -- it's very easy for a 2 to become a 7 in the scanning process.

☎ **If we give it to you, we'll have to give it to everyone.**

So what's the problem with this? We actually get his excuse a lot -- and it's not a reasonable excuse according to our public records act.

☎ **That uses proprietary software**

You don't want to software, you want the data. If they don't know how to copy to a file or PRINT to a file, find out who their vendor is.

One of the latest excuses I got:

☎ **We don't mind giving you a few records, we just don't want to give you the whole database**

To which the judge asked: You mean if they wanted ONE record from ONE person it would be okay?

Their answer: depends on the person.

Costs: What should you expect to pay for data?

media
disk $2 per disk
9track tape $20 per tape

In many cases they can e-mail you the database.

Processing: find out how much the data processing person makes and how long it will take.

Have them itemize the costs

Make sure it's not something they do already

Ask them to do a second backup copy

For more information:

Freedom of Information Center
573/882-4856
University of Missouri School of Journalism
Columbia, MO 65211
jourke@muccmail.missouri.edu

Reporters Committee for Freedom of the Press
800 18th St., NW
Washington, DC 20006
202/466-6313

IRE's Access Committee
IRE: 573-882-2042
Jennifer LaFleur and Jim Neff Co-chairs

* see attachment from Access committe on benchmark costs

7

Talk Like a Nerd to Win Data and Influence Agencies

Megabytes, density, ASCII. Mention these terms and many of us turn pale. But knowing a few computer terms can not only increase your chances for a hot date with a local systems engineer, it also can unlock the door to a wealth of information. Suppose you were interested in investigating federal contracts granted to Bay Area companies. You could examine summary reports some government office created, but you might want to look at the records themselves. Imagine sorting through thousands of documents. Making sense of such information would take a VERY long time. Looking at that same information on a computer makes it easier, if you know what you're doing.

Knowing what to do isn't instantaneous. You can't just pull up to your computer and run a query on a relational database at will. But a beginning to the process is to think about data --eek! Instead of counting sheep as you drift off at night picture tiny flying diskettes and you'll be on your way.

The information in your computer is a bunch of little on and off switches that form different patterns to stand for different things. Each switch is called a BIT -- that stands for Binary DigIT

It's how the computer stores information
- 1 or 0
- off or on
- yes or no

For example: 1= 1
               2 = 10
               3 = 11

Information is just like your old secret decoder ring where a certain pattern would stand for a certain letter. For example, J51 might have stood for the letter A. Your computer does the same thing with different patterns of bits. In order to have enough combinations for all the possible letters, numbers and symbols, your computer needs eight slots -- that group of eight on/off switches is called a byte. That secret decoder language that your personal computer uses is called ASCII. Another type of language is EBCDIC -- that takes more work to translater, but if it's all you can get say "Yes, thank you -- I'll take it."

And, for those of you who want to sound especially techno-savvy:

    -ASCII (American Standard Code for Information Interchange)
    -EBCDIC (Extended Binary Coded   Decimal Interchange Code)

Now all these bytes may come on several different formats:

    - Diskette -- holds about 1.4 megabytes (1.4 million characters)
    - 9-track tape (PLEASE: don't ask for 8-track tapes unless you're a real fan of 70s music)
    - 3290 tape cart
    - CD-ROM

In addition to the data, you also need some hard-copy information. You must have a record layout -- that's the map to your data -- without your data are useless. You also need to know some format information about the data. Once you've determined if your file is ASCII or EBCDIC (or you may get one of those rare files that is already in a database or spreadsheet). You need to know if it is fixed or delimited. A fixed length file has everything in columns (sort of like what you see if you lookup Census tables on the Internet)

Delimited files have commas (usually) separating each piece of information so your database or spreadsheet program knows which piece to put into the first column, second column and so on... Most of the time these files put "" around fields that are names or cities or Ids because those fields are text -- also called character. That tells your computer that you won't be doing math with that information. Be careful with this one because sometimes information that looks like numbers really should be text. For example, the ZIP code 00102 is text. If you told your computer to make that a number -- it might take off the zeros and give you 102 -- not a known ZIP code that I'm aware of. Numbers are the other comma type of information -- those are the things that you do want to do math with such as money or an age. You'll also run into other types such as date, but a good rule of thumb is to keep those fields as text -- you can always change them later.

Here's an example of a delimited file:

"Sumony","Will","tap dancer",100,"CA101"
"Goode","Art S.","student",250,"CA102"
"Roosevelt","Eleanor","housewife",100,"CA101"

The other comma type of file is fixed length. That means that everthing is in columns. So instead of looking for commas, your computer knows that the first field is in slot 1 - 15 and the second field is in slote 16 - 26. This is the way mainfame 9-track tapes usually come. Here's an example:

```
         1         2         3         4         5         6         7
1234567890123456789012345678901234567890123456789012345678901234567890

Sumony         Will      tap dancer      100  CA101
Goode          Art S.    student         250  CA102
Roosevelt      Eleanor   housewife       100  CA101
```

As you can probably tell, the delimited file takes up a lot less space. It's also the easiest form for you to get in your computer. SOOOOO... if you have a choice of format choose ASCII, comma delimited. (These also might be called CSV files (stands for comma-separated values).

Especially, if you have a fixed file, you have to be able to tell your computer which slot every field can be found in -- that's where the record layout comes in handy. Here's the record layout for the file shown above:

Other words for the same things
Most government agencies really won't care that you use Access or Paradox as you databases manager -- they won't give you terms that fit your software. Here's some comma terms:

    Comma-delimited   CSV   comma-quote delimited
         *note: these files also can be delimited with other things such as [or ^ --
but it means the same thing.

    Fixed-length   fixed format   fixed record   standard data format

    Text   character   alpha-numeric

    Numeric: sometimes comes in several flavors, such as integer, double or currency.

Key Terms:

Record: a related group of information in a database (the rows)
Field: each piece of information you have for a record (the columns)

If you get 9-track tapes here are some terms you'll need:

Labelling: this is information put on the tape that gives the basic info about the tape   such as record size, block size and whether it's ASCII or EBCDIC.

Block: a group of records
Blocksize: the number of records in a block

Density: this is how much stuff is squeezed onto an inch of tape. Typically this is either 6250 bytes per in or 1600 bytes per inch. (most tape drives read both these days)

Format: you need to know if the data is ASCII or EBCDIC

Record length: the number of bytes (or characters) in each record

The number of records

Data types:

Character or text: a series of letters and/or numbers that are not computed
Numeric: values that you might do math with
Date: dd/mm/yy or some other standard format
Currency: money

Resources in California and elsewhere

California Journalism Online: http://ccnet4.ccnet.com/CSNE/. This site contains links to
* The California Society of Newspaper Editors
* The Associated Press News Executives Council
* The California Newspaper Publishers Association, which includes the organization's legislation report.
* California First Amendment Coalition (CFAC). The CFAC site contains reports on legislation and case law, which are useful resources when fighting for data access.

Freedom of Information Center: http://www.missouri.edu/~foiwww/
Established in 1958, the center now has a collection of more than 1 million articles and documents about access to information. The Center is open 8 a.m. to 5 p.m. CST weekdays. Call (573)882-4856 or contact Kathleen Edwards, Manager at jourke@muccmail.missouri.edu

Reporters Committee for Freedom of the Press: http://www.rcfp.org
The Reporters Committee for Freedom of the Press is a nonprofit organization dedicated to providing free legal help to reporters and news organizations.

Society of Professional Journalists, Northern California Chapter's Freedom of Information Committee: Co-chairs are: Dan Borenstein of the Contra Costa Times 510-943-8248 and Elizabeth Pritzker at the First Amendment Project in Oakland 510-208-7744.

Listservs

California Freedom of Information Listserv: CAL-FOI. This list, sponsored by the Northern California Chapter of SPJ's Freedom of Information Committee is a forum for discussion of freedom of information, public records and First Amendment issues in California.
Subscribe to maiser@rosebud.berkeley.edu with the message: subscribe cal-foi [your name]

Calgovinfo: This list from Computer Professionals for Social Responsibility provides a forum for discussion of access and computerization of government records in California.
· Subscribe to calgovinfo@cpsr.gov

Books

Paper Trails: A Guide to Public Records In California. Second Edition. Stephen Levine and Barbara T. Newcombe. Center for Investigative Reporting and California Newspaper Publishers Association. Call CIR at 415-543-1200 or CNPA at 916-443-5991.

California: A Public Records Primer and Investigator's Handbook. Don Ray. ENG Press Burbank. Call 818-THE-NEWS for more information.