

Introduction to Databases (MS Access)

A HANDS-ON APPROACH

Purpose

Provide hands-on experience with standard database and reporting operations.

Prerequisites

- You are familiar with saving files to folders/directories
- You should have completed the “Gentle Introduction to Information Systems” worksheet in Appendix, especially the section “Database Preparation”

Meta-awareness: Why does the worksheet have you answer these questions by hand? Computers are only tools to do logical work and calculations more quickly and/or with more data. If you can do it by hand with repeatable steps simple, you most likely can do it with a computer. Additionally, it is important to be sure you know what the answers should be before creating the computer solution. That way, you can run a test situation and confirm that you coded it properly.

Why should you care about databases?

- Jobs that require you to look at large amounts of data: If you ever have to work in a library or market research company, you will feel more comfortable with the tools used there.
- Just as in the spreadsheet exercise, we are setting up the database with only a handful of data that could have easily been analyzed by hand. However, this is just preparation and testing for when we may have to analyze a larger set of data. Testing small will give us more confidence in our queries and reports when our data set increases to 100s and 1000s of students.
- Many commercial packages use a database underneath: personal planners, address books, inventory programs. If you understand how the database needs the data arranged and represented, you can better understand what these applications can and can't do.

Notes:

- If you are already familiar with Access, you may just want to skim this lab. However, if you are new to the PC computer or this program, this lab will step you through some common activities.
- In this lab, `courier font` will be used to indicate text appearing on the computer screen. Any text that you should type in will be underlined.
- This worksheet refers to an Office 2000 application. Newer applications should be similar.

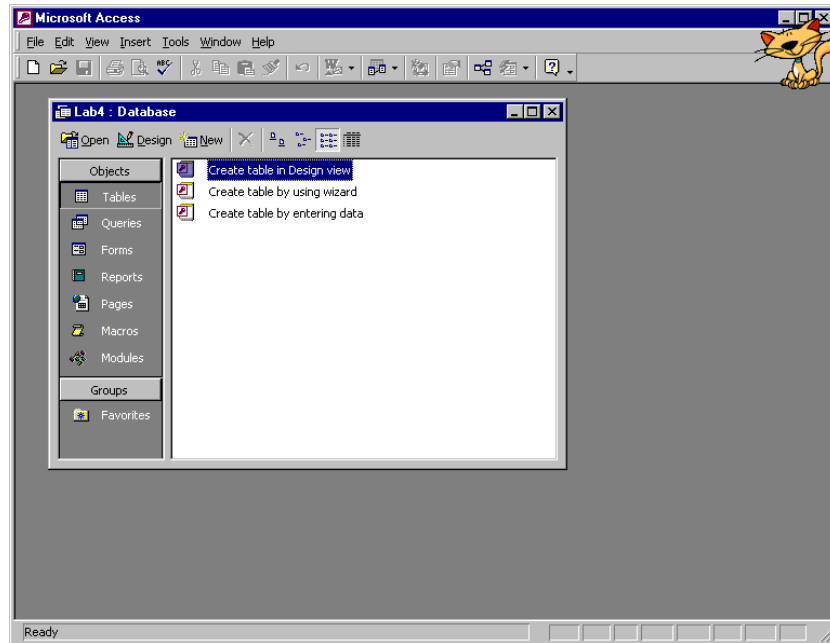
OBJECTIVE

Many applications use databases. We will set up a simple database and query and then show you how to use the Microsoft Access wizards to set up reports. Some programs don't provide such help in setting up reports, but hopefully, you will see some patterns in how data is retrieved.

This activity will guide you through creating a database for the class. It will also introduce you to Wizards, which guide you through a particular activity. When you are done with this project, you should be able to use the database to print a report that answers the questions posed in class.

PREPARATION

The data you collected regarding the students in class will be used in this worksheet.



STARTING MICROSOFT ACCESS

1. **Select** "Start" from the bottom of the screen.
2. **Select** "Programs", then "Microsoft Access". *Microsoft Access starts.*

You may be presented with a start box. Close it for right now by pressing the "Cancel" button.

CREATE A NEW DATABASE

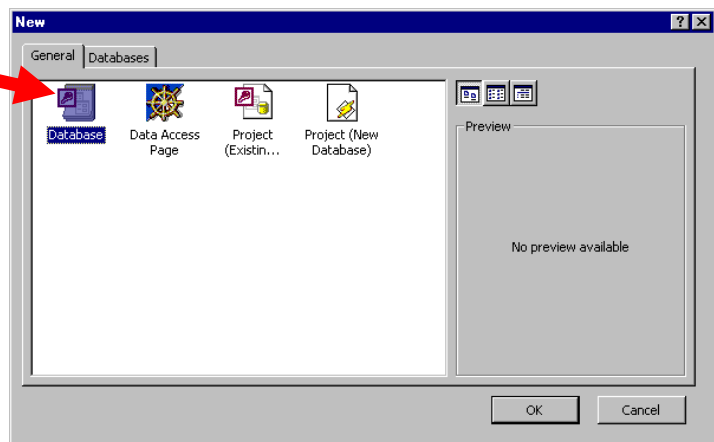
In Access, a database is the combination of Tables, Queries, Forms, and Reports. In other database applications, a database might simply be a single table.

1. **Select** "File" from the top menu bar to create a new database. *The "New" dialog box will appear.*
 - If "New" is not listed, select the double arrow at the bottom of the drop box. When the "New" option appears, select it.
2. **Select** "Database" from General tab.
3. **Click** the "OK" button.

Access next asks you to define the file name for the database. It will present you with a mini-Explorer.

4. **Select** a working drive and directory
5. **Type** Classmates in the File Name box.
6. **Click** the Create button.

The screen show above will appear. This represents the database that contains your class's profile.

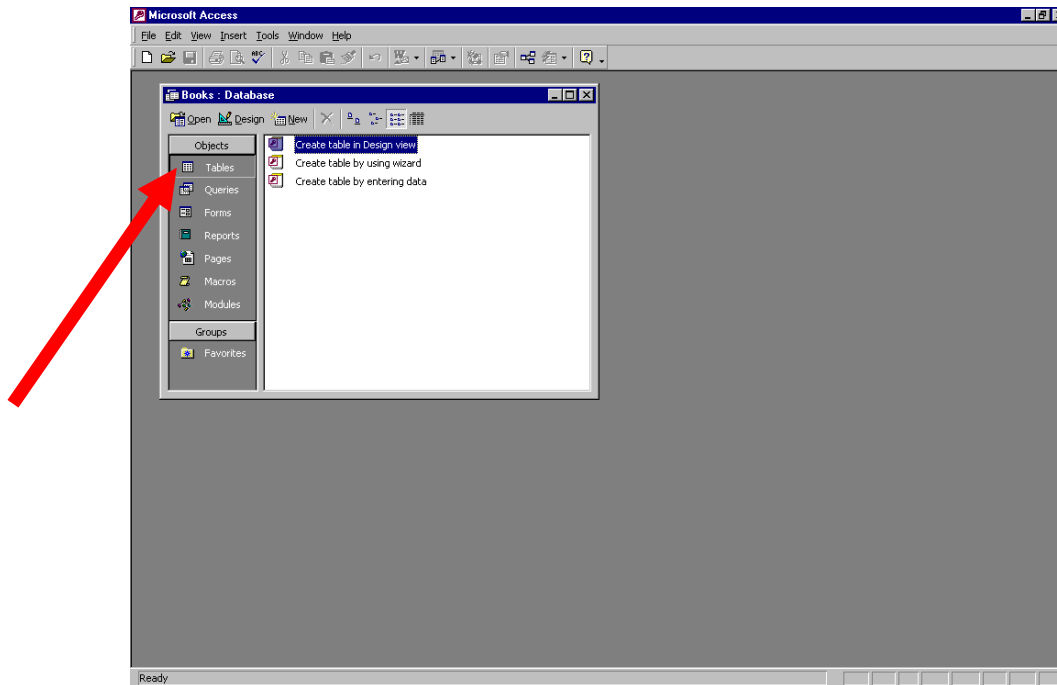


CREATING THE TABLE

Create a table of students

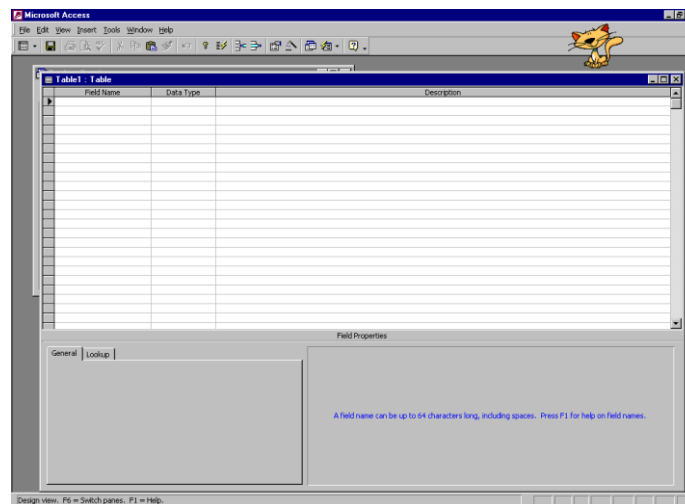
Tables contain the actual data. Think of each student filling out her information on index cards. The tables are collections of information cards.

We will use the “Design view” creation option. This is the most basic method and most databases have a similar creation option.



1. **Select** the “Tables” option under the “Objects” heading (above).
2. **Double-click** the item “Create table in Design view” in the right part of the window.

The following table appears.



3. **Fill** the grid with the following values:
 - a. The data types should have been determined in your preparation for this lab.
 - b. Writing Field Names without spaces will make the analysis easier (as in TotalChildren and FavoriteFood)

Field Name	Data Type	Description
Student	Text	Put whatever description you want in here
Major	Text	
Female	Yes/No	
Sisters	Number	
Brothers	Number	
TotalChildren	Number	
Moved	Yes/No	
FavoriteFood	Text	

4. **Select** "File" from the top menu bar, then "Save" to save the table
 - Again if "Save" is not listed, select the double arrow, then "Save" when it appears on the list.

The program will ask you for the table name.

5. **Type** Students as the table's name.

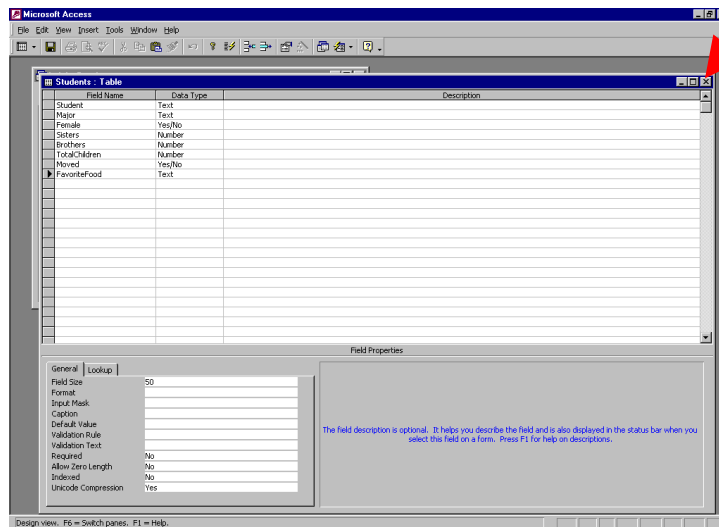
The following dialog will appear. Access wants you to define a primary key.



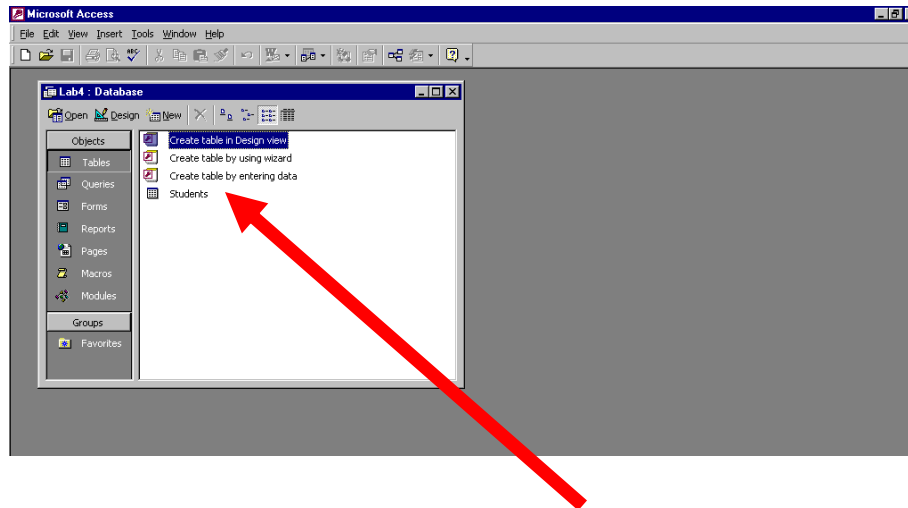
- A **primary key** is like the call number on a book (for example, the Dewey Decimal index or Library of Congress index). It uniquely identifies the book. No other book has the same call number. Primary keys speed the sorting and searching process, but since we are using such small numbers of data, we will choose to not assign one for now.

6. **Click** "No" to skip defining the primary key for this activity
7. **Close** the table definition screen by clicking the x box in the upper right hand corner of the table definition screen, **not** the Microsoft Access screen (right)

You are returned to the Database screen.



Enter Student Data

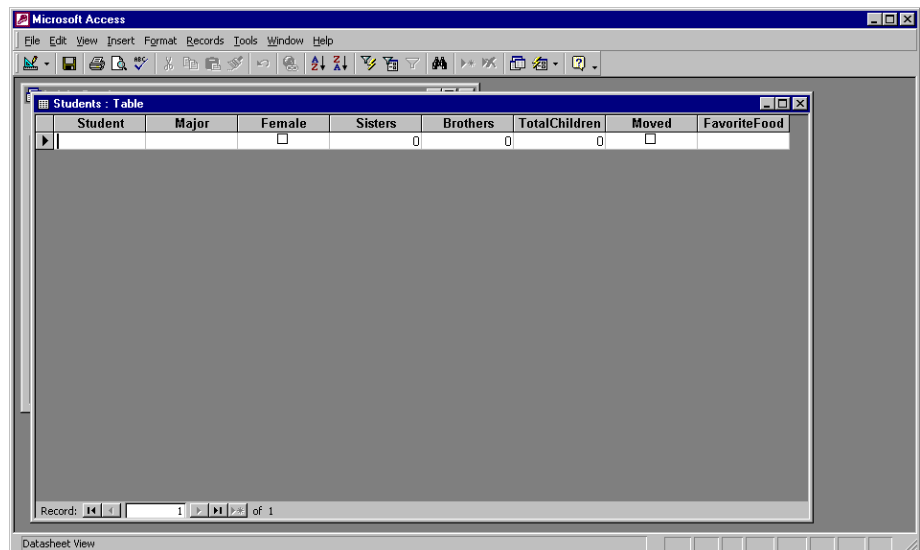


1. **Double-click** on the Students table in the Database screen to open it (above).
2. **Enter** the data you collected from the play in the data view (below). For example,

Student	= Yvonne
Major	= Mechanical Engineering
Female	= Yes (checked)
Sisters	= 1
Brothers	= 0
TotalChildren	= (leave blank)
Moved	= Yes (checked)
FavoriteFood	= Pizza

Print the Table Contents

1. **Select** "File" from the top menu bar, then "Print".
2. **Click** the "OK" button in the print dialog box.
What do you notice about the print out?
What is printed in each column?
3. **Adjust** the table column widths as you did in Excel. Does this change the information printed out?
4. **Close** the table when you are done by clicking the x box in the upper right hand corner of the table screen.
 - If you changed the layout of the screen, you will be prompted "Do you want to save changes to the layout of table 'Students?'" Click Yes if you want to save these layout changes. Click No if you don't care to save them.

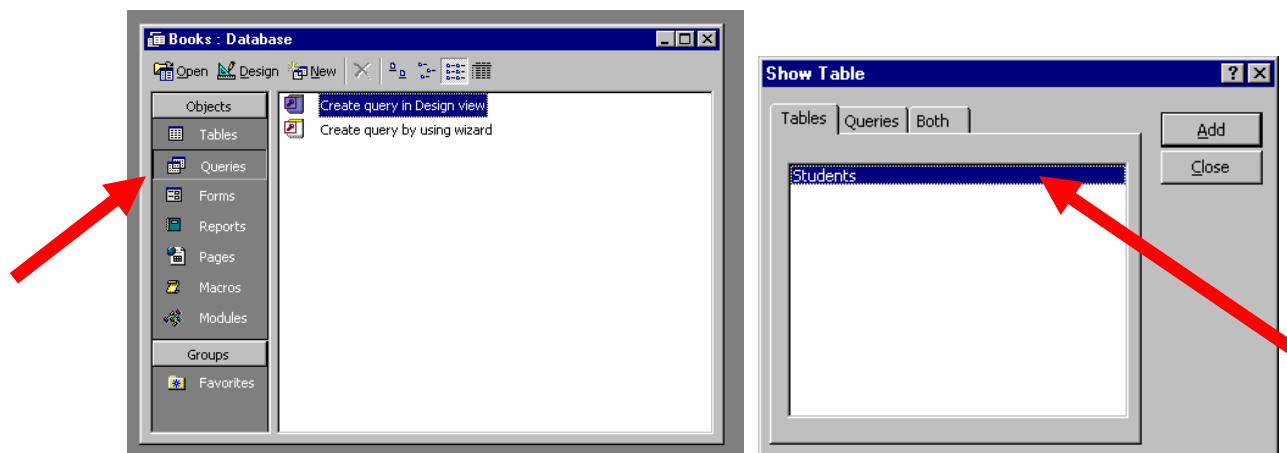


CREATING THE QUERY

A Selection Query describes to the database which field information we want to view and which particular students we want to retrieve.

We will create a query to view the students in your profile. We will use the "Design view" creation option. This is the most basic method and most databases have a similar creation option.

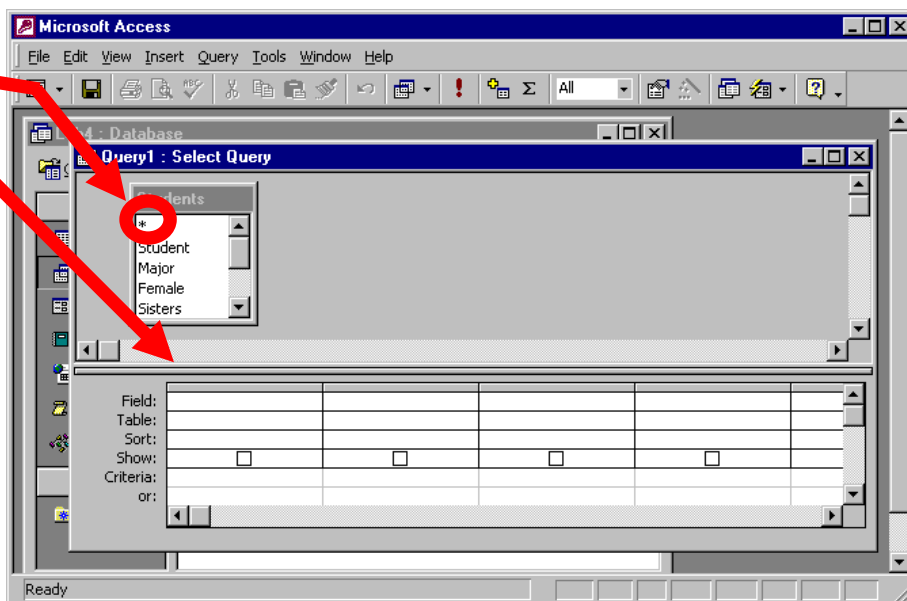
CREATE SELECTION QUERY

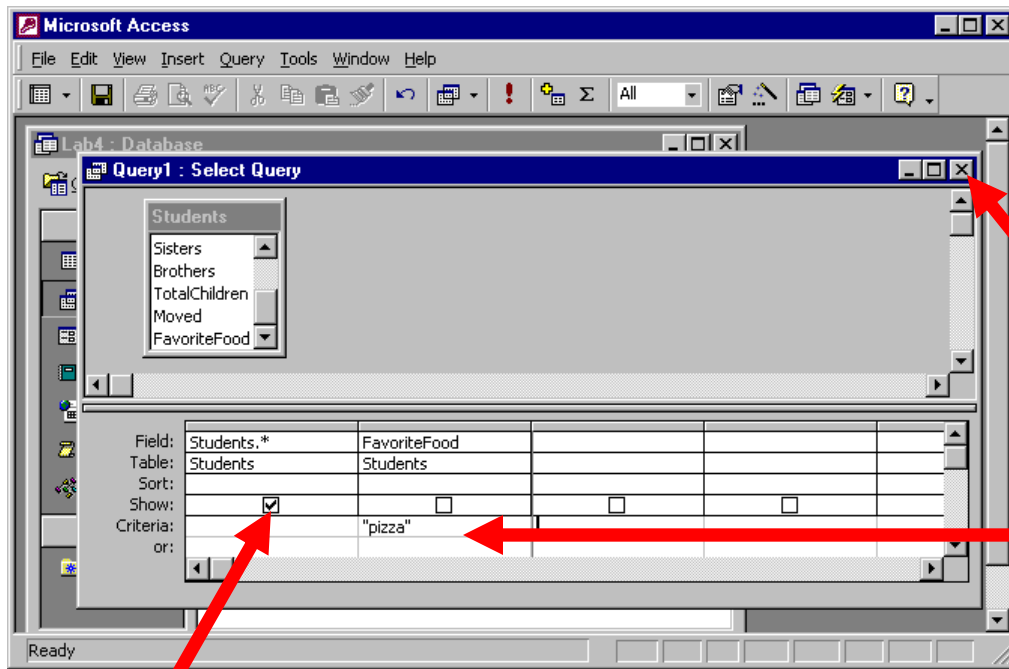


1. **Make sure** you are in the Database screen
2. **Select** the "Query" object from the left column (above-left)
3. **Click** the "Create query in Design view" item in the right side. The "Show Table" appears.
4. **Select** "Students" table from the list of Tables (above-right).
5. **Click** the "Add" button.
6. **Click** "Close".

You will see the screen shown below. The following instructions guide you through selecting the data by Favorite Food.

7. **Drag** the * element from the list and drop it in the first box in the grid.
8. **Select** "View" from the top menu, then "Datasheet View" to see what data was selected. *You should see the data from the Students table.*
9. **Select** "View" from the top menu, then "Design View" to return to the design screen.
10. **Drag** the "Favorite Food" field to the bottom grid.





11. **Uncheck** the Show box. *This prevents the field from appearing in the Datasheet View.*
12. **Type** "Pizza" in the Criteria box under the FavoriteFood field. *This selects all Students that have pizza as a favorite food.*
13. **Switch** to the Datasheet View at this time. *You will see only the students who have pizza as a favorite food*

Save the Selection Query

1. **Select** "File" from the top menu bar, then "Save".
2. Type Students who like pizza as the name of the query.

Close the Selection Query

1. **Click** the x on the upper right corner of the query screen when you are done (above).

You are returned to the Database screen.

CREATING THE REPORT

Unfortunately, not all database applications have the same reporting methods. In fact, defining reports can get rather complicated. As a result, many applications create Wizards to help the user (you) create formatted reports quickly and easily.

Create a Report

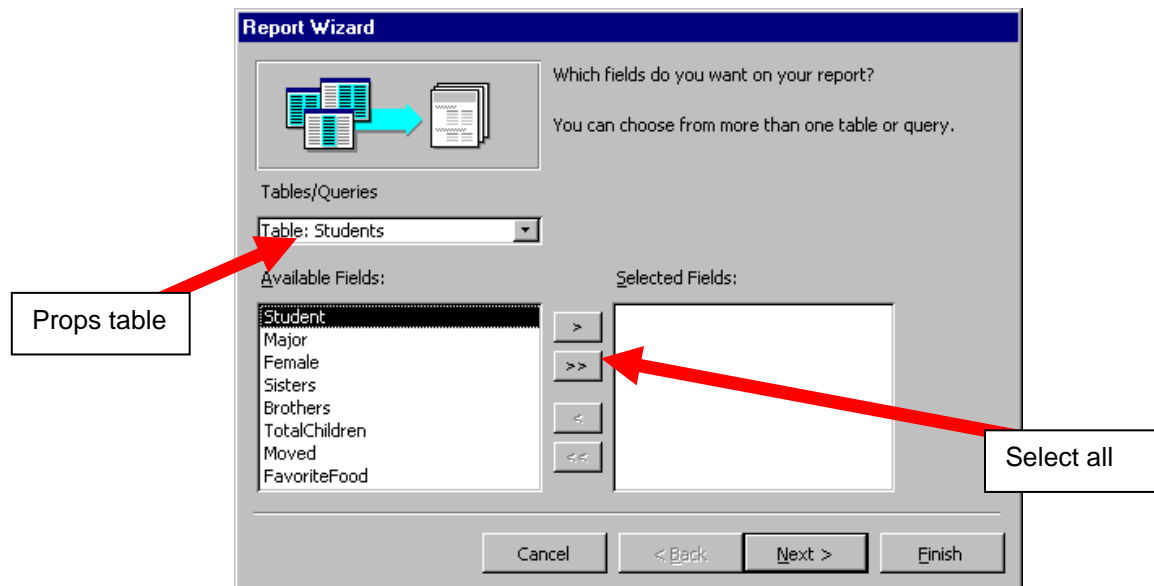
Before creating a report, we need to have an idea of how we want the report to look, which depends on the questions we want to answer. This section steps you through creating a report that will answer the question “What are the favorite foods of the class? Who likes each food?”

We expect the report to look something like:

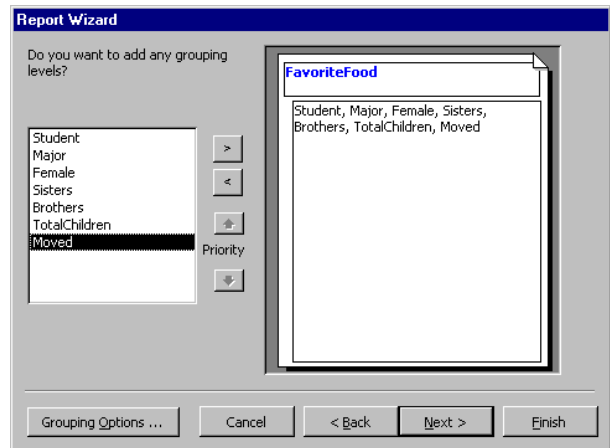
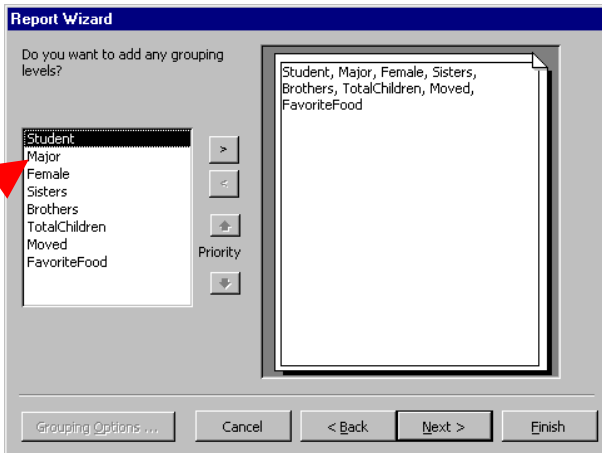
Pizza
 Yvonne
Cookies
 Troy
Pasta
 Helen
...

This means that we want to group together the students with the same favorite food.

The Report Wizard can help us with this.



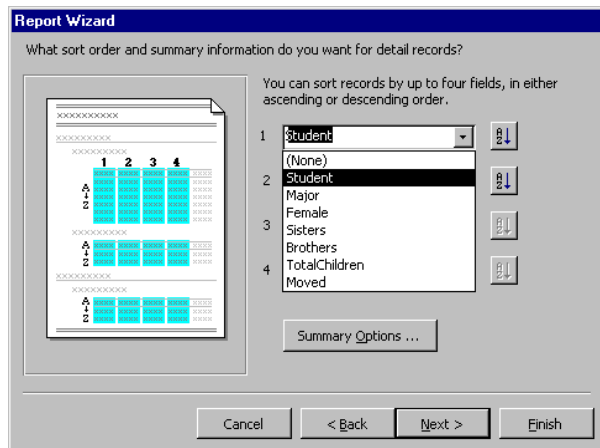
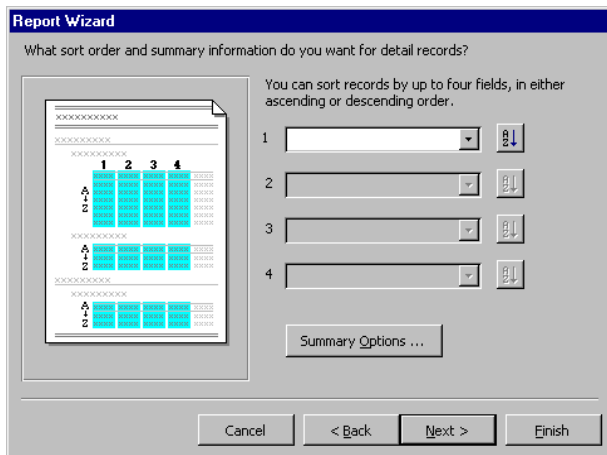
1. **Select** the “Report” item from the Database screen.
2. **Click** the item labeled “Create report by using wizard”. *The first screen of the wizard appears (above)*
3. **Select** the Students table.
4. **Select** all the fields in the Students table by clicking the >> button.
5. **Click** the Next button to go to the next screen.



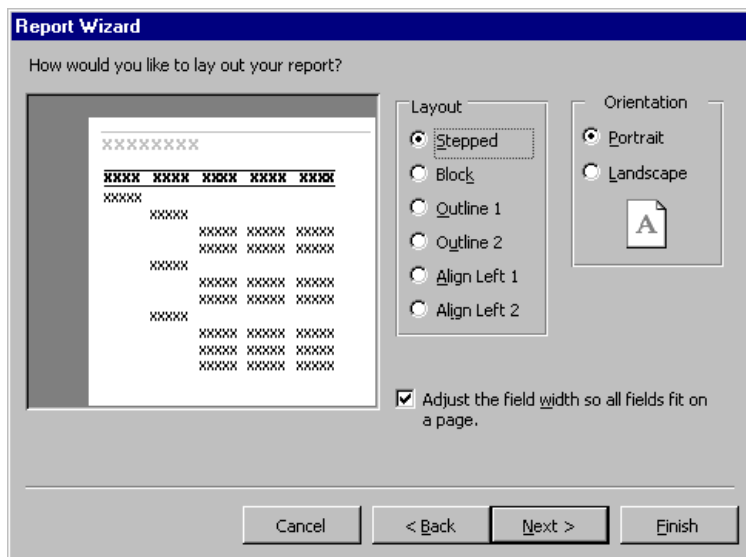
6. **Select** the FavoriteFood field name to group the data by FavoriteFood (above-left).

When you click the > button, the picture on the right should adjust to show Favorite Food as a group heading (above-right).

7. **Click** the Next button to go to the next screen.



Let's sort the students in each favorite food. This is done by using the sort wizard given above.



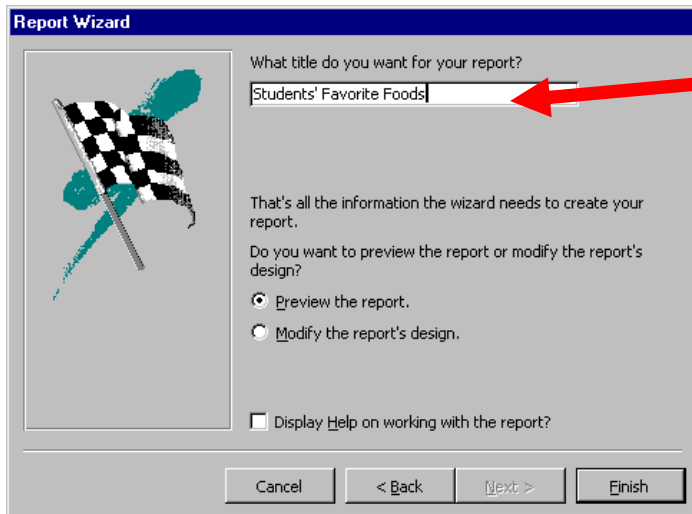
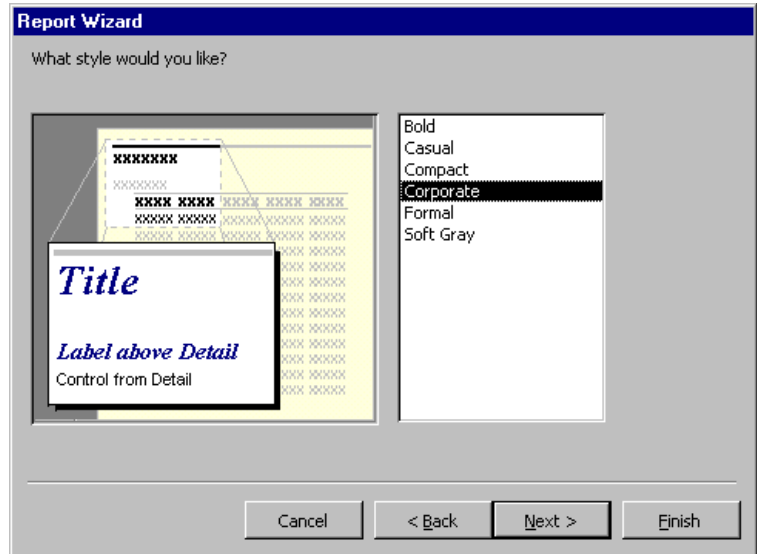
8. **Select** the Student field (above).

9. **Click** the Next button.

10. **Select** a layout (your choice). This example (right) has selected "Stepped".

11. **Click** the Next button.

12. **Select** a style (your choice). This example (below-right) has selected "Corporate".
13. **Click** the Next button.



14. **Type** Students' Favorite Foods as the title for your report. *The report will be saved under this name as well.*
15. **Make sure** the Preview report button is selected.
16. **Click** the Finish button to preview the report (below).

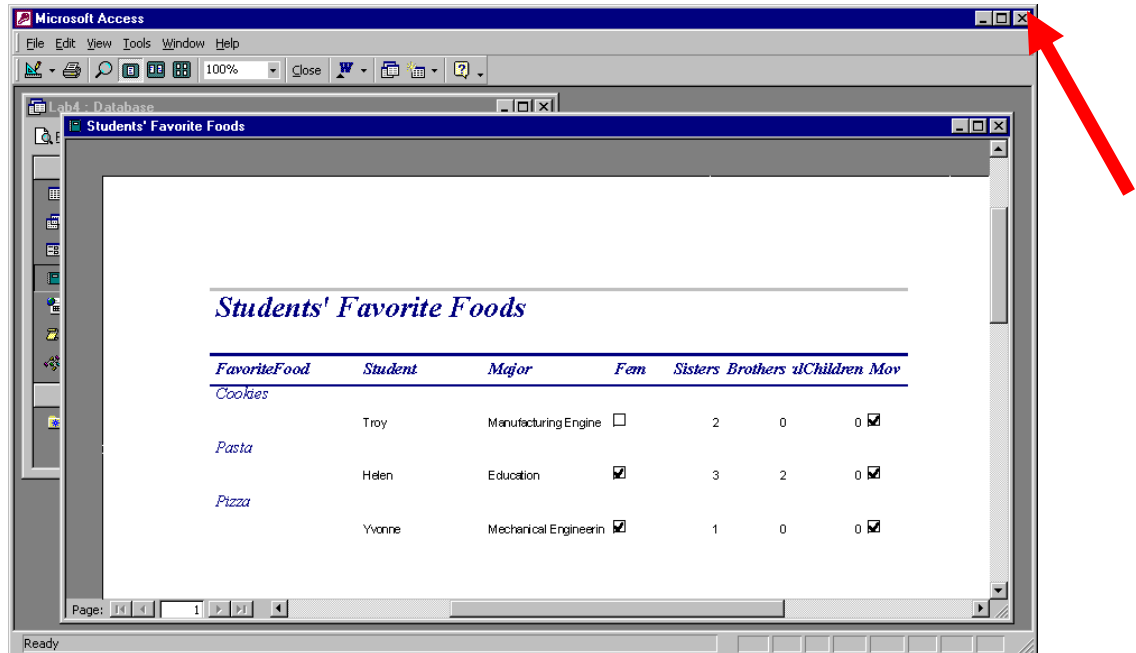
You can modify the report in Design View by clicking the design button. You will be able to resize or change labels. If you tell the instructor what you want to change, she can help you modify the design appropriately.



Save the Report

The report has already been saved, but you can use "File" from the top menu then "Save" to save the report again.

Print the Report



These instructions assume the report preview is still showing:

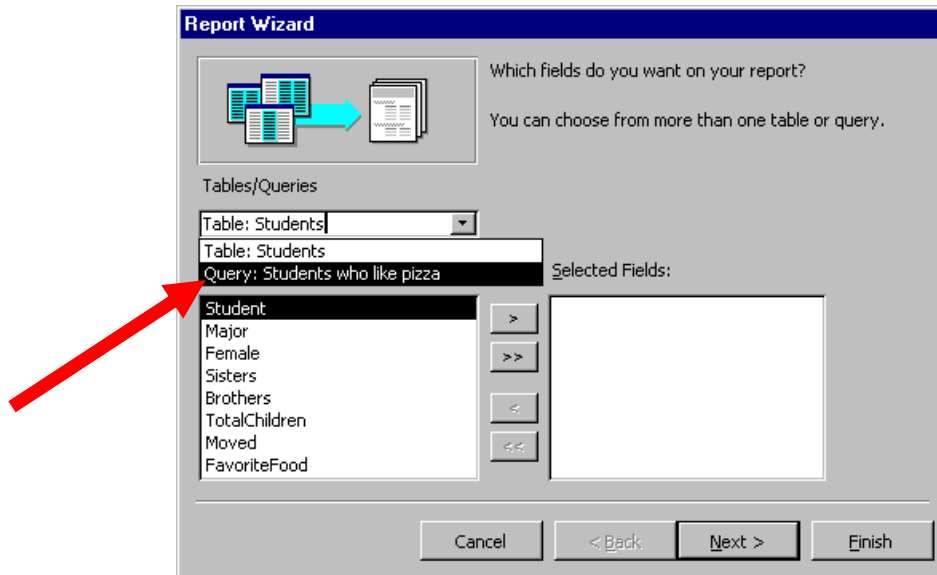
1. **Select** "File" from the top menu bar, then "Print".
2. **Select** "OK" to print the report.
3. **Close** the report by clicking the x in the upper right corner (above).

You are returned to the Database screen.

REPORTING QUERY RESULTS

Create a Report from a Query

The above section took you through creating a report from a table. You can also create a report from a query.



1. **Select** "Report" from the Object list on the Database screen.
2. **Click** the item labeled "Create report by using wizard". *The first screen of the wizard appears.*
3. **Select** the "Students who like pizza" query from the "Tables/Queries" box (above).
4. **Select** all the fields in the query except for the Favorite Food field. *This field will be "pizza" for all items in the query, so why bother printing it?*
5. **Click** the Next button to go to the next screen.

The following directions are exactly the same as in the above section.

6. **Click** the Next button to go to the next screen. We won't group by anything right now.
7. **Select** the Student field to sort by Student name.
8. **Click** Next to go to the next screen.
9. **Select** a layout. Since there is no grouping, just pick "tabular" to see the data listed in rows
10. **Click** the Next button to go to the next screen.
11. **Select** a style (your choice).
12. **Click** the Next button to go to the next screen.

Save the report under a different name

13. **Type** Students who like pizza. The database will also save the report under this name as well.
14. **Make sure** the Preview report button is selected.
15. **Click** the Finish button to preview the report.

AGGREGATE QUERIES

Let's now consider the following question: "What are the favorite foods and how many students like them?"

This report should look something like this:

<u>Favorite Food</u>	<u>Number</u>
Pizza	1
Cookies	1
Pasta	1

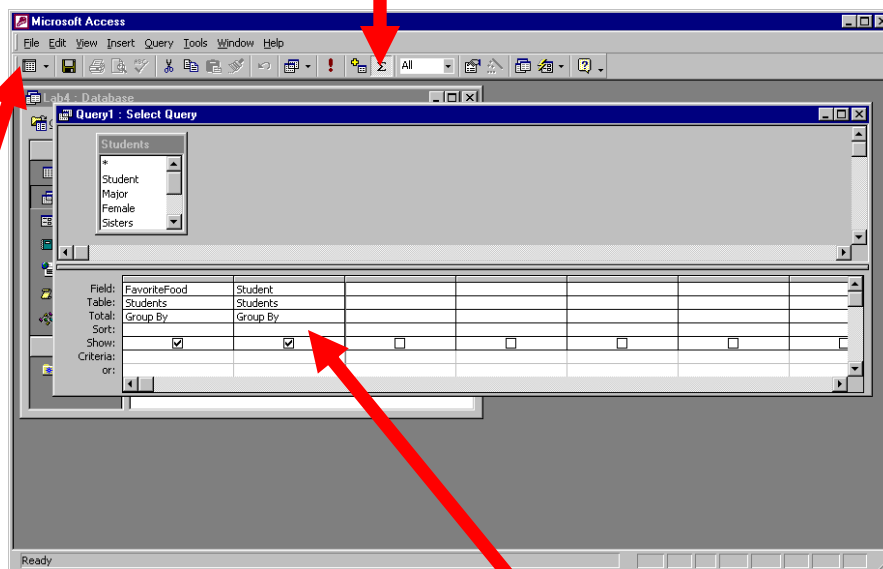
How can we get this grouping information?

Let's go back to our query tool:

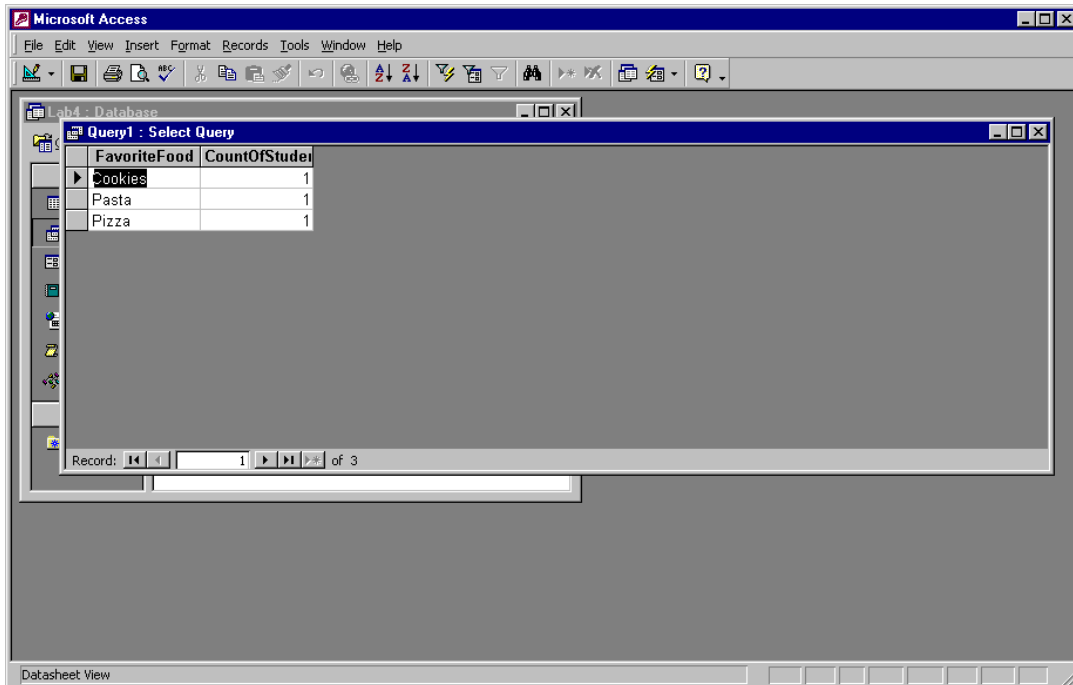
1. **Make sure** you are in the Database screen
2. **Select** the "Query" object from the left column (above-left)
3. **Click** the "Create query in Design view" item in the right side. The "Show Table" appears.
4. **Select** "Students" table from the list of Tables (above-right).
5. **Click** the "Add" button.
6. **Click** "Close".

Now we will start building our summary information

7. **Select** "FavoriteFood" field form the field list and drag it to the first column in the query.
8. **Select** "Student" field from the field list and drag it to the second column. We are using this field since it is non-blank for all valid students.
9. **Click** the Σ button located in the toolbar. This allows you to group the common values.



10. **Change** "Group By" to "Count" in the "Total:" row for the Student field. Leave the FavoriteFood field alone since we want the data grouped by this.
11. **Click** the datasheet view at the top left to see what the data looks like. It should show the information as we wanted it to (next page)



Self-test question

- How would you modify the query to display the Favorite foods in descending order, from Pizza to pasta?
- How would you modify the query to display the foods in order of popularity? The datasheet should look something like:
 - Pizza 5
 - Pasta 3
 - Cookies 1
- How would you create a query that answers the question “What is the average number of sisters for the education majors?” The datasheet should look something like:
 - Education 2.5
- How would you create a query that answers the question “Which students have sisters and how many do each of them have?” The datasheet should look something like:
 - Helen 3
 - Yvonne 1
 - Troy 2

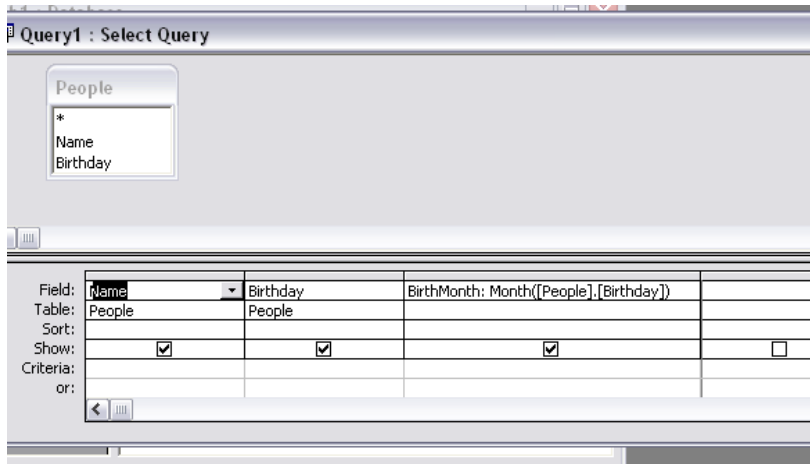
Hint:

- Remember that you can sort the datasheet using field values
- Take a look at the other aggregate functions available such as Avg and Where

Other helpful tips

Expression

You can process a field value to create a different expression which can be used in a query. For example, the screen below shows how the Month() function can be used on the birthdate field to show the month (as a number). Notice how the Expression is defined by *Expression: function([Table].[Field])*. The data view on the right shows how the month is reported by the expression.



Name	Birthday	BirthMonth
Yvonne	6/19/1969	6
Celeste	7/30/1980	7
Mom	6/16/1946	6
*		

Now that you know how Month() works, look it up in Help and see how it describes its usage.

Other useful functions: - look up what they do and usage in Help.

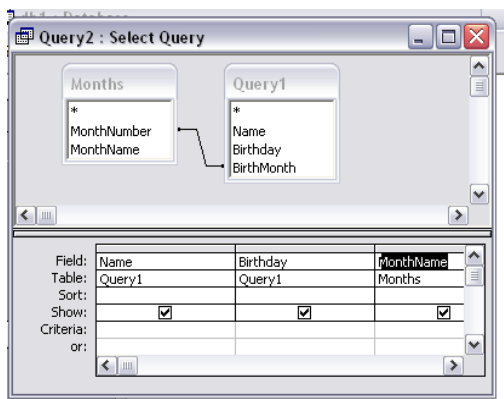
- Day(), Year(), Today(), Left(), Lower(), Concatenate()
- Mid(), Trim(), Substitute()

Joining Tables

It is possible to set up a table to “look up” values. For example, if you wanted to display months, you would need to

1. Set up a table of month numbers and month names (this is the lookup table) (right)
2. Create a query that joins the MonthNumber field from the Months table to the BirthMonth field from the Query (below left). Do this by pressing down on the field and dragging your mouse to the matching field. A link line will be made between the two fields. Be careful, order matters.
3. The result will show the name corresponding to the month (below right)

MonthNumber	MonthName
1	January
2	February
3	March
4	April
5	May
6	June
7	July
8	August
9	September
10	October
11	November
12	December



Name	Birthday	MonthName
Yvonne	6/19/1969	June
Celeste	7/30/1980	July
Mom	6/16/1946	June

Your results don't match Query2? Can you change your join so it does?

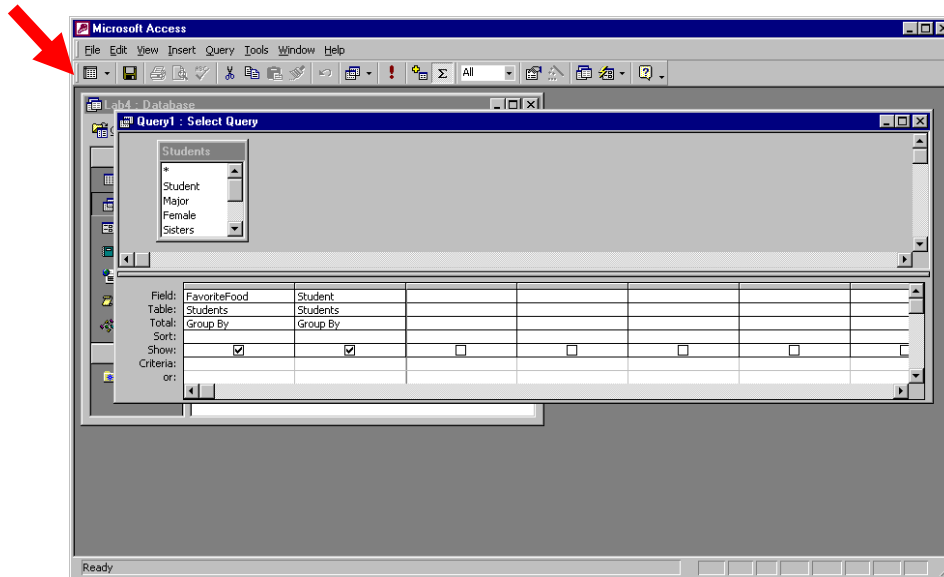
SQL and the query concepts

SQL, sometimes called “S-Q-L” or “See-quill”, stands for Structured Query Language. Some consider it an example of 4th generation programming language.

- 1st generation was machine language, 1’s and 0’s
- 2nd generation was assembly language. It used words, but a very limited set.
- 3rd generate was general programming languages that were easier to read that assembly and permitted implementation of common elements like input, output, variables, arrays, conditionals, and loops/iterative as well as arithmetic and logical operations such as AND, OR, and NOT. This included languages like COBOL, FORTRAN, C, APL

4th generation languages were more specialized, designed for specific functions. With the advent of databases being used widely in industries (e.g. airline reservations, banking), the ability to quickly define logical queries resulted in the specialized languages, the most popular being SQL for databases.

You can see the SQL Microsoft uses by clicking the dataset button in the menu bar and selecting “SQL”. This is the syntax used by many Microsoft products: Access, Excel, Visual Basic, etc. It differs slightly from other SQL applications, but the general elements of SQL are similar no matter what application you use. We will focus on this general aspect so you can dissect the SQL that is generated by the Query Wizard.



1. Open the Students who like pizza query
2. Toggle to SQL mode
3. There are two main keywords to find:
 - a. SELECT
 - b. FROM
 - c. WHERE
4. Notice the format:
 - a. SELECT *FieldName*
 - b. FROM *TableName*
 - c. WHERE *FieldName* = “*value*”
5. Different SQL systems differ in the syntax used in the names. For example: In Access
 - a. If the Field name has a space in it, the name will be surrounded by []
 - i. SELECT [E-mail Address], Company
 - b. If the Table name has a space in it, the name will be surrounded by []
 - i. FROM [My Class]

- c. The format of the WHERE depends on the field's value
 - i. "value" for text
 - ii. value for number
 - iii. #mm/dd/yyyy# for dates
 - iv. "Yes" or "No" for Boolean

Here is a list of conditional and logical operator used in MS Access

https://support.office.com/en-us/article/table-of-operators-e1bc04d5-8b76-429f-a252-e9223117d6bd#_toc272228349

More detail on these syntax aspects for MS Access can be found here:

<https://support.office.com/en-us/article/access-sql-basic-concepts-vocabulary-and-syntax-444d0303-cde1-424e-9a74-e8dc3e460671>

TEST YOURSELF

Test yourself by seeing how these SQL styles are similar or different from MS Access's SQL format:

- MySQL: <https://www.tutorialspoint.com/mysql/mysql-select-query.htm>
- Generic SQL: https://www.w3schools.com/sql/sql_select.asp

Test yourself by looking at the join query. Things to note:

- This will have two tables listed
- Tables are including as part of the field name (especially needed if the field names are the same in the two tables). Format is *[Table Name].[Field Name]*
- There is a JOIN keyword. This indicates that the tables are joined.
- There will be an ON keyword. This indicates which field in the first table joins with which field in the other table.
- Access defaults to INNER JOINS
 - Inner join selects only the rows that match in both tables. In the join example, only the months that were in the `Students who like pizza` query
 - Change INNER to LEFT. LEFT joins always selects the values in the left query (Months in this example). If there isn't a match in the right table, fields related to the right table will remain blank. It can help to put in a field from the right table to distinguish the rows that have a match from the ones that don't
 - Change LEFT to RIGHT. This is the reverse of LEFT. It will show all the records in the right table and show only the matching values in the left table.

Here is more information on join types:

<https://support.office.com/en-us/article/join-tables-and-queries-3f5838bd-24a0-4832-9bc1-07061a1478f6>

Here is more information on how to create shorthands for table names

<https://support.office.com/en-us/article/access-sql-from-clause-2a7d031c-6912-4e8c-a644-590940ddaed1>

Test yourself by create a query with an aggregate and looking at the SQL. Things to note:

- SELECT will have an `count()` expression
- WHERE is related to the grouping condition
- There is a GROUP BY keyword

Test yourself by adding a condition to this aggregate query (e.g. where Count of students is more than 3). Things to note:

- HAVING will replace WHERE

Here is more information related to the HAVING keyword:

<https://www.w3resource.com/sql/aggregate-functions/count-having.php>

Here is more about more complex conditions with aggregate queries:

<https://docs.microsoft.com/en-us/sql/ssms/visual-db-tools/use-having-and-where-clauses-in-the-same-query-visual-database-tools?view=sql-server-2017>

Watch the Data Model

It's easy to see that joining tables can make a lot of rows. When querying large tables, a trick is to first query the large table with conditions related to that table's attributes and create a temporary table. Then a query is set up with the smaller table and the temporary one.

Here is information on how to create tables with the MS Access UI:

<https://support.office.com/en-us/article/create-a-table-query-96424f9e-82fd-411e-aca4-e21ad0a94f1b>

Here is the syntax to create a table:

<https://support.office.com/en-us/article/select-into-statement-74d958dc-4b35-4b97-a7f7-72087e036d9c>

QUESTIONS TO ANSWER

With this data, we want to answer the following questions about the class:

- How many students are in the class?
- How many are Education majors?
- What % of students is female?
- What % of students has moved at least one time in their life?
- How many students like pizza?
- What are the favorite foods of the class? Who likes each food?
- What is the average number of sisters for the education majors?
- Which students have sisters and how many do each of them have??

Spreadsheet Preparation

Spreadsheets are good for counting and calculating.

What are the steps you would use to find the answer the following questions using only the data sheet on page 1 and pencil and paper?

How many students are in the class?	
How many are Education majors?	
What % of students is female?	
What % of students has moved at least one time in their life?	
How many students like pizza?	

Database Preparation

Databases are good for reporting grouping details into meaningful categories.

What are the steps you would take to answer these questions using only the data sheet on page 1 and pencil and paper? How would you arrange your information to clearly answer the question?

- What are the favorite foods of the class? Who likes each food?

- What is the average number of sisters for the education majors?

- Which students have sisters and how many do each of them have?

DATA TYPE ANALYSIS

Before you set up the columns in your database, determine the data type for each field (column). Check the appropriate data type for each field.

Field Name	Text	Number	Yes/No	Date/Time
Student				
Major				
Female				
Sisters				
Brothers				
TotalChildren				
Moved				
FavoriteFood				