

**Massachusetts Institute of Technology
Department of Urban Studies and Planning**



**11.204: Planning, Communication, and Digital Media
Fall 2004**

Recitation 6: Using Access to Query Multiple Data Sets

1. What is database?

- Remember stack cards in libraries?
- Management of students, or employees...
- Gift registry for wedding
- My literature database

Various definitions of a database (<http://www.google.com/search?hl=en&lr=&oi=defmore&q=define:database>):

- an organized body of related information
- A collection of data: part numbers, product codes, customer information, etc. It usually refers to data organized and stored on a computer that can be searched and retrieved by a computer program.
- A collection of information organized and presented to serve a specific purpose. (A telephone book is a common database.) A computerized database is an updated, organized file of machine readable information that is rapidly searched and retrieved by computer.
- A collection of related information about a subject organized in a useful manner that provides a base or foundation for procedures such as retrieving information, drawing conclusions, and making decisions.
- A shared collection of logically related data, designed to meet the information needs of multiple users in an organization. The term database is often erroneously referred to as a synonym for a "database management system (DBMS)". They are not equivalent. A database is a store of data that describe entities and the relationships between the entities. A database management system is the software mechanism for managing that data.
- A collection of related types of data in a single file or set of files for sorting, analysing, and reporting.

2. What are the advantages of computer-based database?

Imagine what you did (or would have done) in the libraries before computer-based systems (like OPAC) were introduced.

- Less work (searching for data/matching two different sets of data)
- Cross-referencing
- Accessibility/Portability

3. Why do we need Access?

- In lab 4, from where is the population density field? (Look in ArcMap)
 - Area--ESRI shape file
 - Population--Census 2000
- How do we use Census 2000?
 - Free from the Census Bureau: [Method 1](#) or [Method 2](#)
 - Private vendors (like [Geolytics](#)) sell Census CDs, but they are expensive.

3.1 What's the difference between MS Excel and MS Access?

- Excel--non-structured or free structured; flat file
- Access--strictly structured based on a general standard

4. Input -- System (Core Function) -- Output --User Control

Core Function of Access

- Data collection, manipulation, integration--generate new datasets with more information or more specific information

User Control Interface

- MS Access Interface and Elements

Data come from different resources in different formats.

- Excel, Text, Lotus, dBase, Oracle, StarOffice...
- Population (Census Bureau), Crime Record (Police Department), Zoning and Parcels (BRA), Traffic Flow (Transportation Statistics Bureau), Fire Loss (Fire Department), Tax Data (Internal Revenue Service)

Data can be used for different purpose and further analysis.

- Regression or Cluster analysis (SPSS), Make graphs and charts (Excel), Create Maps (ArcView, ArcGIS), Web Publish (HTML)...
-

5. MS Access Operation

5.1. Framework of MS-Access database

- Table: database part. operates like Excel.
 - Design View: Definition of columns
 - Primary Key: ID# for each record (e.g. call number at the library)
 - Datasheet View: Real data
 - Composed of records (rows), columns (fields), and cells.
- Query: tool (1) to select specific data under criteria, and (2) to integrate multiple database.
- Form: tool/interface to view the contents of the database
- Report: tool to print out the contents of the database

5.2 Creating Our Own Database by Importing Data (Input)

Several ways to create a database--importing and designing the structure, then inputting data

Create a new empty database. Save the database.

Import Tables

- Which worksheet to use
- First row contains column headings
- Store the data in a new table or an existing table
- Field Name, Type, Index or not, Setup a primary key
 - Type:
 1. Text

Use for text or combinations of text and numbers, such as addresses, or for numbers that do not require calculations, such as phone numbers, part numbers, or postal codes. Stores up to 255 characters. The **FieldSize** property controls the maximum number of characters that can be entered.

2. Memo

Use for lengthy text and numbers, such as notes or descriptions. Stores up to 65,536 characters.

3. Number

Use for data to be included in mathematical calculations, except calculations involving money (use Currency type). Stores 1, 2, 4, or 8 bytes; stores 16 bytes for Replication ID (GUID). The **FieldSize** property defines the specific Number type.

4. Date/Time

Use for dates and times. Stores 8 bytes.

5. Currency

Use for currency values and to prevent rounding off during calculations. Stores 8 bytes.

6. AutoNumber

Use for unique sequential (incrementing by 1) or random numbers that are automatically inserted when a record is added. Stores 4 bytes; stores 16 bytes for Replication ID (GUID).

7. Yes/No

Use for data that can be only one of two possible values, such as Yes/No, True/False, On/Off. Null values are not allowed. Stores 1 bit.

8. OLE (Object Linking and Embedding) Object

Use for OLE objects (such as Microsoft Word documents, Microsoft Excel spreadsheets, pictures, sounds, or other binary data) that were created in other programs using the OLE protocol.

Stores up to 1 gigabyte (limited by disk space).

9. Hyperlink

Use for hyperlinks. A hyperlink can be a UNC path or a URL. Stores up to 64,000 characters.

o Index

An index helps Microsoft Access find and sort records faster. Access uses indexes in a table as you use an index in a book: to find data, it looks up the location of the data in the index. You can create indexes based on a single field or on multiple fields. Multiple-field indexes enable you to distinguish between records in which the first field may have the same value.

You should consider indexing a field if all the following apply:

- The field's data type is Text, Number, Currency, or Date/Time.
- You anticipate searching for values stored in the field.
- You anticipate sorting values in the field.
- You anticipate storing many different values in the field. If many of the values in the field are the same, the index may not significantly speed up queries.

o Primary Key

Each table should include a field or set of fields that uniquely identifies each record stored in the table. This information is called the primary key of the table.

- Name the table

Tables will vary in size based on the number of records that you are importing.

5.3 Working with Tables

Modifying tables

- Changing a Cell Value
- Adding/Deleting Records

Sorting tables

- Simple

Filtering tables

- Simple
- Advanced Filtering with a logical operator

5.4 Select Queries

Simple query based on one table

- Construct a select query that tells you how many census tracts in the Commonwealth of Massachusetts have more Black people than White people.

Query based on table join

- Construct a select query that tells you how many census tracts in the City of Cambridge have more Black people than White people.

Query based on table join, sort, calculation by using query, expression

- Build a select query to calculate the population density (number of people per square mile) for each of tract in the City of Cambridge. Sort the records in ascending order. Which tract has the greatest density?

5.5 Export Query Result (Output)

Excel, dBASE, text, html, Lotus, Word

6. Introduction to Relational Database

6.1 Something behind the queries--SQL (structured query language)

Three-tier model: Client--Web Server--Data Server e.g. Online Application, Online Shopping, Plaza, WebSIS

SQL vs. HTML

SQL vs. Java, C++ Procedural vs. Non-procedural

6.2 One to one, one-many, many-many Relationship

- A one-to-many relationship

A one-to-many relationship is the most common type of relationship. In a one-to-many relationship, a record in Table A can have many matching records in Table B, but a record in Table B has only one matching record in Table A.

Ex: One supplier ... can supply more than one product ... but each product has only one supplier.

- A many-to-many relationship

In a many-to-many relationship, a record in Table A can have many matching records in Table B, and a record in Table B can have many matching records in Table A. This type of relationship is only possible by defining a third table (called a junction table) whose **primary key** consists of two fields — the **foreign keys** from both Tables A and B. A many-to-many relationship is really two one-to-many relationships with a third table. For example, the Orders table and the Products table have a many-to-many relationship that's defined by creating two one-to-many relationships to the Order Details table. One order can have many products, and each product can appear on many orders.

Ex: One order can have many products ... and each product can appear on many orders.

- A one-to-one relationship

In a one-to-one relationship, each record in Table A can have only one matching record in Table B, and each record in Table B can have only one matching record in Table A. This type of relationship is not common, because most information related in this way would be in one table. You might use a one-to-one relationship to divide a table with many fields, to isolate part of a table for security reasons, or to store information that applies only to a subset of the main table. For example, you might want to create a table to track employees participating in a fundraising soccer game. Each soccer player in the Soccer Players table has one matching record in the Employees table.

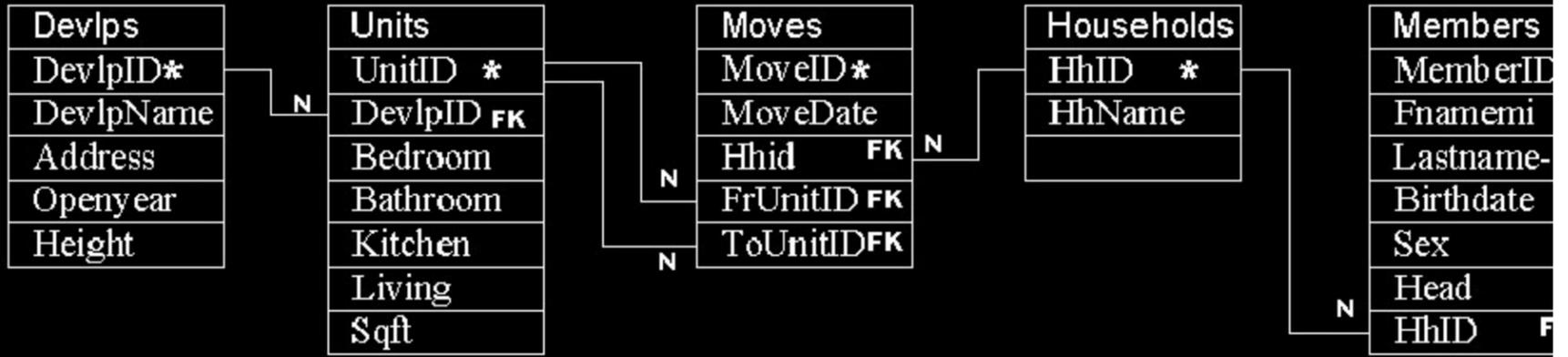
Ex: Each soccer player has one matching record in the Employees table.

- About defining relationships

The kind of relationship that Microsoft Access creates depends on how the related fields are defined:

- A one-to-many relationship is created if only one of the related fields is a **primary key** or has a **unique index**.
- A one-to-one relationship is created if both of the related fields are primary keys or have unique indexes.
- A many-to-many relationship is really two one-to-many relationships with a third table whose primary key consists of two fields — the **foreign keys** from the two other tables.

Public Housing Database Schema



Legend

- N** Many Side
- *** Primary Key
- FK** Foreign Key

7. Database Vendors and Products

VENDORS	PRODUCTS	COMMENTS
ORACLE	ORACLE	Giant, Mainstream
IBM	DB2	
Microsoft	SQL Server	Powerful
MySQL AB	MySQL	Open Source, Free, Plaza
Filemaker	Filemaker Pro	User friendly