Introduction to Databases

Data organized yields information

What is a Database

- A collection of pieces of information (data) which is stored in a specific structure in such a way that there is minimum repetition of data and normally serving multiple purposes.
- Database software allows data to be stored, updated, organized, rearranged, and retrieved.
  - Ability to query & generate reports
  - Ensure data validity through reduced redundancy, relationships, and access controls.
  - Customized to allow easy to use interfaces.
    - A person using doesn't need to know how to set up databases, reports or forms.

Data → Information

- Data alone provides little information
  - Example data: Keoki, Leilani, Emma, Jeremy
- Data which is organized yields information.
  - Example information: ICS101 students scoring 100% on test are Keoki, Leilani, Emma, Jeremy
- Database programs allow data to be organized in numerous different ways which provides different types of information.

Have you ever used a database?

They’re Everywhere!

Uses of Databases

- Course listings
- Student records
- Mailing lists
- Movie Listing
- Online Stores
- iTunes
- And much, much more

Building Blocks of a database

- Field - the smallest piece of meaningful data
  - Each field is named and defined.
  - Example of fields: first name, last name, address.
- Record - group of fields relating to one entity
  - Example: A student record with above fields
- Table - a group of similar records containing information about a number of entities.
Database vs Information base

- A database has a specific structure with fields, records, tables.
- An information base does not store information in a specific structure.
  - Search engines can be considered an information base because they search entire websites and not specific fields.

Building Blocks of a Database

<table>
<thead>
<tr>
<th>Field</th>
<th>Students Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student ID</td>
<td>Sam</td>
</tr>
<tr>
<td>First Name</td>
<td>Leilani</td>
</tr>
<tr>
<td>Last Name</td>
<td>Suzuki</td>
</tr>
<tr>
<td>Phone</td>
<td>274-6765</td>
</tr>
<tr>
<td>Address</td>
<td>10 Decade Ave.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Record</th>
<th>Students Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>12345</td>
<td>Sam</td>
</tr>
<tr>
<td>67890</td>
<td>Leilani</td>
</tr>
<tr>
<td>98765</td>
<td>Wes</td>
</tr>
<tr>
<td>54321</td>
<td>David</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Field</th>
<th>Students Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student ID</td>
<td>IC3101</td>
</tr>
<tr>
<td>Course</td>
<td>A</td>
</tr>
<tr>
<td>Grade</td>
<td>F2011</td>
</tr>
<tr>
<td>Semester</td>
<td></td>
</tr>
</tbody>
</table>

Defining Fields

- Fields can be defined with:
  - Name
  - Type
  - Size
  - Text – number of characters.
  - Numbers – type integer, byte.
  - Formats
  - Other options

- Field types:
  - Text – store text and numbers which don’t need to be calculated
  - Name, address, phone number
  - Numbers – store as numbers only numbers that are calculated.
  - Salary, quantity, credits, grade points, Date and Time
  - Birthday, sales date, enrollment date
  - Yes/No (Logic) – Any field which only has yes or no (true or false)
    - Represented by a check box
  - Memo – long text fields
    - Comments, reviews
  - More types are available

Tables Store Data

- Tables store data in records which are made up of fields
  - All records in a table have the same fields defined
  - Each record in a table refers to one entity
  - A group of records is called a table
- Databases allow us to see the data in any number of different ways but these views of the data do not store the data. Data exists only in the tables.
  - Update the data in a form and the data in the table is updated

Views

- Views allow data to be seen in different ways, but don’t store data
- Views do not STORE data, but display the data that is already stored in a defined manner
- Types of views:
  - Query – displays particular information from a database.
  - Form generally designed as an input device for a database. Information is shown on screen.
    - May be based on a query
  - Report – Presents data in an organized manner
    - May be based on a query
What are Queries?
- Queries essentially provide answers to questions. Queries allow data to be selected, displaying only the data that is needed subset of total records
- parts of a record
- Select Queries select data according to specified criteria

Forms & Reports
- Input of data.
  - Forms are designed for input of data
  - Reports do not allow input
- Primary record views
  - Forms generally display one primary record at a time
  - sometimes with multiple related records
  - Reports generally display multiple records

Databases vs Spreadsheets
- Databases
  - allow better report generation with updated data.
  - help ensure data validity through reduced redundancy, relationships, and access controls.
  - can be customized to allow user friendly interfaces.
    - End user doesn’t necessarily need to know how to set up databases, reports or forms.
- Spreadsheets allow for easier numeric calculations

Relational Databases
- Linking Tables

What Are Relational Databases
- A relational database allows us to tie together (or relate) the tables within our database allowing us to create queries, reports and forms from multiple tables.

Reducing redundancy
- Relational databases store data in a format that reduces redundancy.
  - We want to reduce redundancy because it affects the integrity of a database.
  - EXAMPLE: If a person's address is stored in 3 or 4 tables and the person moves, there is an integrity problem if not all tables are updated.
    - In relational databases, the address is stored once and related to other tables, so changes need to be updated in only one place
  - Access is an example of a relational database
Why Use Table Relationships?

- Reduces redundancy
- Ensures integrity of data
- Allows more flexibility in data use
  - Reports and forms based on from various related tables
    - Looks like one record even though it is stored in separate records.
  - Information can be pulled from various tables to meet different applications or uses of the database.

Types of Relationships

- One-to-one
  - For one entity, there is only one possible piece of data for a given attribute
    - Person to first name
    - Book to title
  - Usually defined in one table
- One-to-many
  - For one entity, there are more than one possible pieces of data for a given attribute
    - Student to courses
    - Book title to copies in library
  - Usually define in separate tables

One-to-One Relationships

- For each student, there is only one possible data for each of these fields
  - Student ID
  - First Name
  - Last Name
  - Home Phone
  - Street Address
  - City
  - State
  - Zip
  - Financial aid status

One-to-Many Relationships

- For each student, there are any number of Courses Completed - One student may have 0, 1, 2, ... 100 or more courses
- One student may have any number of student work hours. Some students have no work records as they are not employed as a student worker.

How Do Relationships Work?

- Use of key fields - All tables containing the key can be related
  - For one table the key is primary and a unique key is assigned to each record.
  - For other tables the key is secondary and is used as a link to the table where the key is primary.
    - Table with secondary keys may also have a primary key that uniquely identifies that table's record.
  - Sometimes a table is created to establish links between two tables (linking table)

What is Database Integrity?

- Integrity in data bases means that the information is accurate and consistent.
- Relational databases help insure integrity by having data stored once and tying the data together.
- Good forms help insure integrity
  - If data coming in is accurate, information being output will be accurate.
Reports & Forms
- Can pull information from various related tables
  - Information looks like it is all one record even though it is stored in separate records.
  - Information can be pulled from various tables to meet different applications or uses of the database
- Queries can also have multiple tables and can assist with creation of reports and forms.

Relationships in Access
- Defining Key
  - Primary table must have primary key defined.
    - Define key by selecting field then click on the key icon.
  - In the secondary table the secondary key is defined exactly the same as in the primary table but does not as primary key
    - Don't click on the key icon
- Establish relationships
- Database Tools - Relationships
- Add tables
- Drag key field from primary table to key in secondary table.
- Enforce referential integrity

Access Wizards
Recommendations from teacher about Access Wizards
- Queries – don’t use the query wizard – does not allow criteria to be input
- Do Use Wizards in Access to create Forms & Reports

Database Queries
Finding the information needed

What is a query
- Queries assist with finding the right data for the task - Display only the data needed
  - only certain records
  - only particular fields
  - or both
- Queries generally need a criteria to select the correct data
- Queries can also
  - Create calculated fields
  - temporary data generated from a calculation
  - Allow certain ACTION on the database

Queries into Reports & Queries
- Reports generally need only certain data rather than all the data
  - Base reports on the queries
  - Only the data from the query will show in the report.
- Queries can be used to create more sophisticated reports and forms
  - Calculated fields
  - Data from multiple tables
  - Data meeting certain criteria.
- Create a query first, then base the form or report on the query.
### Types of Queries

- **Select queries**
  - Select data based on a criteria
  - Create calculated fields
- **Parameter queries**
  - Same as a select query, but the criteria is entered whenever the query is run
  - Different criteria can be entered each time
- **Cross-tab queries**
  - Cross-tabulates using multiple criteria
- **Action queries**
  - Update queries
  - Update records quickly.
  - Sometimes using a calculation
  - Global updates based on criteria
- **Deletion queries**
  - Delete sets of records based on criteria
- **Append queries**
  - Add records from one table into another table

### Criteria

- Queries select data based on a specified criteria.
- Criteria are specified for particular fields.
- Criteria specified must match the type of field.
- Text field must have text criteria (numbers can be part of a text field, too).
- Number fields must have number criteria
  - Leave off formatting such as dollar signs and commas.
- Date fields must have date criteria

### Criteria Qualifiers

<table>
<thead>
<tr>
<th>To specify</th>
<th>Use Symbols</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equal to (same as)</td>
<td>=</td>
</tr>
<tr>
<td>Greater than (after)</td>
<td>&gt;</td>
</tr>
<tr>
<td>Less than (before)</td>
<td>&lt;</td>
</tr>
<tr>
<td>Greater than or equal to (on or after)</td>
<td>&gt;=</td>
</tr>
<tr>
<td>Less than or equal to (on or before)</td>
<td>&lt;=</td>
</tr>
<tr>
<td>Not equal to</td>
<td>&lt;&gt;</td>
</tr>
</tbody>
</table>

### Boolean Criteria

- Boolean logic can also be used
  - **AND**
  - **OR**
  - **NOT**

In the next software assignment you will be asked to enter a code for this lesson. The code is Criteria. Make note of it.

### Example Criteria

- Find all records where location is Kailua
  - Kailua
- Find all records where location is not Kailua
  - <>Kailua
- Find all records where salary is equal to or over $20,000
  - >20000
- **Logic Criteria**
  - Yes - finds all records where this field is checked
  - No - finds all records which are not checked

### Multiple Criteria

- Multiple Criteria can be used with Boolean Logic
  - Find all records where location is Kailua OR Kaneohe OR Waimanalo
  - Find all records where locations is Kailua AND salary is over $20,000
  - Location field criteria: Kailua
  - Salary field criteria: >20000
  - Criteria must be placed on the same line or becomes an OR and returns more items
    - all records where location is Kailua and
    - All records where salary is >20000
      - no matter where the location is
### Date Criteria
- Dates are stored as whole numbers that count the days from 1/1/1900.
  - Dates before a specified date are considered less than
  - Dates after a specified date are considered greater than
- Example find people who are less than 30 on January 1, 2011.
  - Birthday is GREATER THAN 1/1/1981
    >1/1/1981

### Logical Thinking Required
- To specify correct criteria, you need to THINK!
  - Which field is needed
  - What information is stored in that field
  - Which qualifier is needed.

### Logical Thinking Example
- Example: Find everyone in the database who lives in Hawaii.
  - Which field would this information be stored?
    - State (or some other similarly named field)
  - What information is stored in that field?
    - Two letter abbreviation -- for Hawaii we query for HI
  - What qualifier?
    - = (equal is the default, so you don’t need to enter it)

### Example cont.
- Be sure to enter criteria on the criteria line under the correct field.

### More Critical Thinking
- Find people who are less than 30 on January 31, 2013.
  - What field is needed?
    - Birthdate
  - What type of information is stored there?
    - dates
    - Query must match the type of data so do not use <30 as that is not a date!
  - THINK! What birthday do people have who are less than 30 on January 31, 2013?
    - Born after 1/31/1983. For dates AFTER use >
    - >1/31/1983

### More Critical Thinking
- Find everyone from out of state
  - Can list all of the other states, but that would be too tedious
  - Critical thinking, if someone is from out of state then they are not in Hawaii
  - Use the Not Criteria
  - <>"HI"
- USE CRITICAL THINKING!!!! You have to translate the request into something the database can compute.
**Calculated Fields**

- Calculated fields allow for calculation of information.
  - For example for a sales order, a calculated field can multiply the quantity times the price to come up with a total for a given item.
  - Instead of storing the total, which would be redundant, we store just the quantity and price, then use a calculated field to provide the total when needed.
- Many functions, such as used in Excel, are also available to help calculate information.

**More on calculated fields**

- Some functions help us deal with text such as displaying a first and last name together.
- In Access, use the Builder tool to create a calculated field within a query.
  - Click on fields to refer to these fields
  - Use functions and mathematical formulas to create calculated field

**Design Mode & Wizards**

- Wizards allow quick generation of tables, queries, forms and reports.
- Design mode allows individual customization of tables, queries, forms and reports.
- Forms and reports generated by wizard can be customized in design mode.
- Instructor recommends
  - Using Wizards for reports and forms
  - Using Design mode for queries The wizard does not ask for criteria and criteria is the backbone of queries!

**Database Integrity**

- Integrity in data bases means that the information is accurate and consistent.
  - Relational databases help insure integrity by having data stored once and tying the data together
  - Being consistent in spelling, word choice help insure integrity also.
    - For instance using Hawaii and the abbreviation HI in the same field in different records causes inconsistency.
    - If we do a query for “Hawaii” we may miss records with the abbreviation HI.
- Accurate Information
  - Remember the computer saying "Garbage in - Garbage out"
    - If data is not accurately entered, the information coming out will not be accurate.
  - Accuracy in input can be increased by
    - Data field properties
    - Form design
    - Action queries
  - Accuracy in output can be increased by
    - Report design
    - Query design
**Action Queries**

- **Append queries**
  - Eliminates reentry of data
  - Allows records from other sources to be added to the database
- **Update queries**
  - Updates records based on a criteria
    - Example: For all records that show dues have not been paid, set member status to inactive
  - Quicker to do and more accurate
    - No records meeting criteria are missed

**Data Field Properties**

- **Lookup fields**
  - Species a limited number of input choices
  - Input can be made through a list of choices
- **Input Masks**
  - Specifies data must be in a certain layout
  - Phone number (999)999-9999
  - Input must be numbers.
  - Parenthesis and dash already included for input
- **Validation**
  - Data must meet certain rules
    - Example <1000

**More Field Properties**

- **Default Values**
  - Use with caution
  - Use only when the initial input data is the same for all records
  - All new members are ranked as Novice
  - or when data changes rarely
  - Year joining an organization only changes once a year
  - Be sure to change default in the new year

**Forms Design**

- Create forms using Form Wizards, but use Design View to customize.
- Clearly label forms
- Include instructions to clarify how to fill out form
- Input aids: Dropdown list, checkbox, radio buttons
- Keep layout organized in a familiar way
  - Addresses laid out in the same way one would normally see an address
  - If based off a paper form arrange similarly

**Action Queries**

- **Append queries**
  - Eliminates reentry of data
  - Allows records from other sources to be added to the database
- **Update queries**
  - Updates records based on a criteria
    - Example: For all records that show dues have not been paid, set member status to inactive
  - Quicker to do and more accurate
    - No records meeting criteria are missed

**More Action Queries**

- **Deletion queries**
  - Deletes records based on a criteria
  - Verify correct records are being selected before running as a deletion query
    - Example: For all members who have been inactive for 5 years, delete the records
  - Use with caution.
  - Be sure you will never need the data again.
  - Before running update or deletion query, enter criteria as a select query to ensure the correct records are found.
Reports Generation

- Usually based on query to have just the information needed
  - Correct query design
  - Test to see if works

- Design reports for ease of use and reading
  - Clearly identify information
    - Appropriate titles, headings, etc.
    - Use new page designation to keep only desired information on each page
  - Use report wizards to ease creation of reports
  - Customize reports in design view

Database Integrity Insured

- If data coming in is accurate, information being output will be accurate.

- User error is always possible
  - Relational databases insure data is input once
  - Ease input to decrease user error

- Design errors can be present
  - Create appropriate queries to derive information
  - Base reports on queries
  - Design both reports and forms using Wizards, but be sure to customize for ease of use