

Bioinformatics: Life Sciences on Your Computer

Database Structures

This is supplemental reading for the database lecture.

Databases

Textbook definition: Computerized archive used to store and organize data in such a way that information can be retrieved easily via a variety of search criteria. Databases allow users to manage and organize data. Usually, multiple parties can use/modify that data, but control over who can modify might be necessary. Databases allow for the distribution of data to multiple users.

Spreadsheets

One of the simplest type of database is the spreadsheet. In most spreadsheets, columns represent **fields**. Each field represents a unique way to identify object, such as name, date of birth, address. Rows represent **records** (also called **entries**). With each record, a user can read characteristics by looking across the line.

fields			
Harridan, R.	97 Starr Ave, Minto	289-9922	records
Harrison, A.	8 Ramsey Rd, Cambelltown	492-5000	
Harrison, G.	123 Blue Jay Way, Liverpool	292-0001	
Hastings, E.	9/12 Longshore, Cambelltown	492-7991	

Discussion topic: Look at the address field above. If you sorted by address, which would come first?

Spreadsheets enable users to sort data. In general, more fields allow for more sorting flexibility. It's often best to separate fields into smaller fields. That allows more varieties of searches.

Harridan	R	97	Starr Ave.	Minto	289-9922
Harrison	A	8	Ramsey Road	Cambelltown	492-5000
Harrison	G	123	Blue Jay Way	Liverpool	292-0001
Hastings	E	9/12	Longshore	Cambelltown	492-7991

Discussion topic: After sorting by town, how many 'secondary sorts' are useful? In what ways is this spreadsheet more useful than the one with fewer fields? What types of organizations would want to sort by address? By phone number?

Field (attribute) types

Composite field: can break to subfields (e.g. address)

Multi-valued field: can have multiple values per record: (e.g. phone numbers)

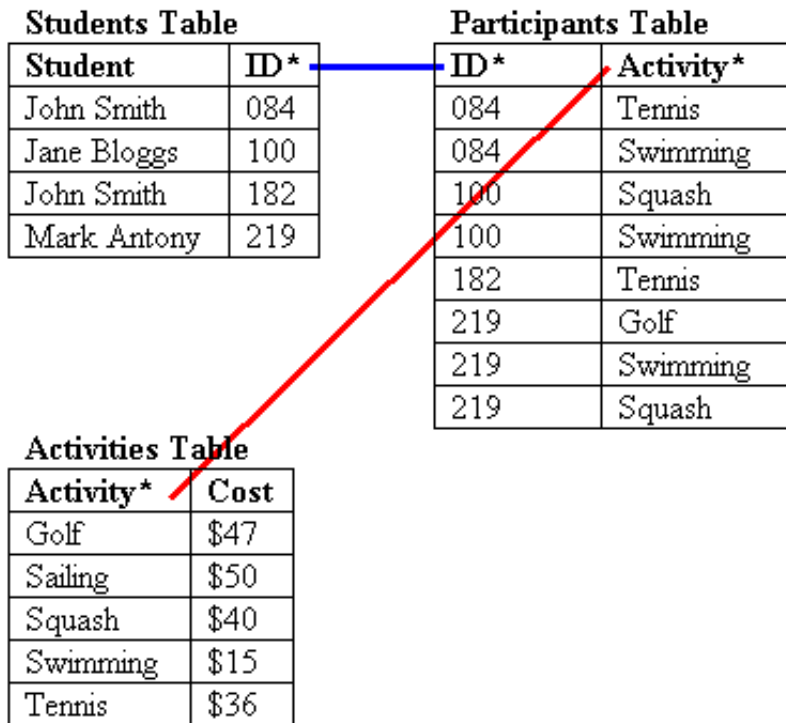
Single-valued field: only one possible entry (SSN, birthdate)

Derived field: can calculate based on other data - Example: Age from current date & birthdate

Base field: field from which other field derived (birthdate)

Relational databases

Many relational databases use SQL (Structured Query Language). Small databases are linked. When you hit one value, you can access info from the linked databases.



System-based Databases

The spreadsheet is the simplest form of database (MS Excel, Lotus). Sometimes it is necessary to vary the field sizes and what type of information goes in each field. The system-based database is usually hosted on a mainframe's platform (e.g. UNIX). A system-based database allows fields of different shapes and sizes. The cursor location biggest limitation because most unix systems require using the arrow keys and not a computer mouse.

Windows-Based Databases

Windows-based databases are a bit more user-friendly than system-based in that a mouse can be employed to quickly move between fields. The computer (network) size can be limiting, but both PCs and Macs no longer have much size limitation. You need not be connected to the internet to access the data.

Discussion topic: Imagine magazine or journal subscription data. If you called to renew a subscription, your data could pop up in a window. What types of custom fields might appear in that window and what type of data would be accepted in those fields?

Web-Based Databases

One of the big advantages of web-based databases is that anyone with an internet connection can

access and modify data. One of the big disadvantages of web-based databases is that anyone with an internet connection can access and modify data.

Obviously security is an issue. The administrative side of the database must be password protection. Here is a screenshot of what the back end of a sports news database looks like. It's old but most news-based databases are similar in concept. An administrator that is watching the newswire can update a news website like CNN from his/her bedroom.

		v Hold down the Control Key to make multiple selections.
Story Title*:		
Story Date*:	01/26/2006 (mm/dd/yyyy)	
Story Article*:	 Display in Activedit	
Image 1 - File:		Upload a File
Image 1 - Caption:		
Image 1 - Author:		
Image 2 - File:		Upload a File
Image 2 - Caption:		
Image 2 - Author:		

When someone registers for something on the internet and provides personal information, that person is creating an entry into a web-based database.