ICT Database
Lesson 1
What is a Database

Database
A file that stores data in an organized fashion so that information can be retrieved from it.
Examples:

- iPod playlist
- Netflix movie list
- Contacts in cell phone

Table – Flatfile Database

- A collection of data organized in rows and columns that can be used to store and manage information
- Work great with small lists of data (information) that is all related.
- A simple way to create a flat file database is to use a spreadsheet.
- Store information in cells created by using columns and rows of data

Spreadsheets - Advantages

- Great for analyzing and sorting related data.
- Easy to learn and use
- Suitable for storing and "crunching" relatively small volumes of numerical data.
- Able to present numerical data in the graphical form to quickly analyze data.

Spreadsheets - Disadvantages

- Require that you enter the same information in multiple places
- Have simplistic sorting and querying capabilities
- Contain only a finite number of records
- Changes to data in the computer memory and are not complete until the file is saved

Data Integrity
The validity of the data
"Garbage in, garbage out."
Creating Tables

- Identify duplicate data
- If duplicate data exists, create a separate table just for that data
- Relate it back to the original table

Relational Database
Consists of multiple tables of information related through common fields
Advantages

- Hold as much as 2 gigabytes of data and are only limited by server operating system.
- Allows for the simultaneous access and query of data by multiple individuals in multiple locations.
- Data is protected against inadvertent corruption so you no longer need to keep redundant data.
- We don't have to manually enter the information reducing time, effort and mistakes.
- Entering information is minimalized which reduces input time, resources and opportunities for human error. This reduces cost and increases data integrity.
- Reduced processing time for large amounts of data.

Disadvantages

- They are more complex and harder to learn and use than spreadsheets

- Designing relational tables can be more difficult and time-consuming and software and hardware costs are higher than a spreadsheet.

Big Data

A term that describes enormous volumes of data that are too huge for regular database programs

- Unlike data in a database, big data is unstructured and unrelated.
- Analysis of big data requires specialized tools.

Database Management Systems

A program used to store, access and manipulate database information

- Microsoft Access is an end-user DBMS that you can use to create and manipulate fairly small and uncomplicated databases
- Oracle, MySQL, IBM's DB2 and Microsoft SQL Server are high-end DBMS programs used to create and manipulate large, complex databases used in large organization.

SQL – Structured Query Language

A sub-language commonly used for developing and managing databases

- SQL pronounced "sequel"
- Used primarily in a database to retrieve, update, insert or remove information
- Sufficiently powerful and can create tables, restructure tables and remove tables, among other very complex tasks

ODBC - Open Database Connectivity

An open standard application programming interface (API)

- Allows us to use the MS Access interface tools to access the DBMS data.
- Using a simplified graphical user interface (GUI) like MS Access to build queries and reports is a flexible way for a relative novice to gain access to a wealth of company data.