Topic 1.1: Data vs. Information

Because the terms Data and Information are important to the design of databases, it is better understand the difference between them.

What is Data?
The term data is used to describe raw facts (not yet processed) about something or some one. Data is raw facts from which the required information is derived.

Example: ROBCOR Company tracks all sales for its two divisions through invoices. Each of the invoices contains raw facts such as these:

Invoice number = 300124,  Invoice date = 12-Jan-2004, and

Sales amount = $125.98

Suppose that ROBCOR's two divisions have generated 1,380,456 and 1,453,907 invoices, respectively, between the first quarter of 1999 and the first quarter of 2004. These millions of invoices the ROBCOR had generated contains raw facts which do not have meaning unless and until they are processed into information.

Now, suppose that for the purpose of making conclusions and/or decisions, the ROBCOR's sales managers want information about sales productivity per employee for each of the two divisions. The generated sales productivity information will reveal meaning of the data exist in the sales invoices. An application program in an information system will generate the required information. Figure 1.1, from the text book, shows a graph of the sales per employee for each of the ROBCOR's two divisions.

**FIGURE 1.1 Sales per Employee for Each of ROBCOR's Two Divisions**
When data is stored electronically in files, it can be used as input for an information system. An information system has programs to process (or transform) data to produce information as an output, see Figure 1. Information reveals meaning of data. For example, students' data values such as ID, Name, Address, Major, and Phone number represent raw facts. Class roll is a list which shows students' ID and Names of those students who are enrolled in particular class (course section).

Let's summarize some key points:

- Data constitute the building blocks of information.
- Information is produced by processing data.
- Information is used to reveal the meaning of data.
- Good, relevant, and timely information is the key to good decision making.
- Good decision making is the key to organizational survival in a global environment.

Timely and useful information requires accurate data. To achieve accurate information, the data must be stored and generated properly. Also, the data must be stored in a format that is easy to access and process. And like any basic resource, the data environment must be managed carefully. Thus, **Data management** is a discipline that focuses on the proper generation, storage, and retrieval of data.

**Concept Check**

What is data?
What is information?
What is the difference between data and information?
List four common types of files?