

Managing Data in a Database

by Sophia



WHAT'S COVERED

In business, an information system's ability to effectively manage data plays a significant role in that organization's ability to manage operations, generate revenue, contain costs, and control risks.

Businesses that seek to gain a competitive advantage must be able to effectively and efficiently store and retrieve the data they collect. In this tutorial, we will discuss data management and how to retrieve stored data.

Our discussion will break down as follows:

1. Data Management

As businesses and organizations have sought to grow and improve products and services, the realization that company data should factor into decision-making has encouraged business leaders to look for data internally and externally. As more data is amassed, the need for systems to effectively manage data is integral. The most important aspect of data management is having access to the data, or being able to retrieve it. Leaders of businesses understand the importance of having access to reliable information at all times, for the purpose of maintaining a competitive advantage. Businesses and organizations must be able to systematically maintain and manage company-related data that can be accessed, retrieved, referenced, and analyzed at any time, so they can make decisions that will ultimately push the business or organization forward.

To the computer, a database looks like one or more files. In order for the data in the database to be read, changed, added, or removed, a software program must access it. Many software applications have this ability: iTunes can read its database to give you a listing of its songs (and play the songs), and your mobile-phone software can interact with your list of contacts. But what about applications to create or manage a database? What software can you use to create a database, change a database's structure, or simply do analysis? That is the purpose of a category of software applications called database management systems (DBMS).

DBMS packages generally provide an interface to view and change the design of the database, create queries, and develop reports. Most of these packages are designed to work with a specific type of database, but generally they are compatible with a wide range of databases. For example, Apache OpenOffice.org Base can be used to create, modify, and analyze databases in open-database (ODB) format. Microsoft's Access is used to work with databases in its own Microsoft Access Database format. Both Access and Base have the ability to read and write to other database formats as well.

Microsoft Access and Open Office Base are examples of personal database-management systems. These systems are primarily used to develop and analyze single-user databases. These databases are not meant to

be shared across a network or the Internet, but are instead installed on a particular device. They work with a single user at a time.

A database that can only be used by a single user at a time is not going to meet the needs of most organizations. As computers have become networked and are now joined worldwide via the Internet, a class of database has emerged that can be accessed by two, 10, or even a million people. These databases are sometimes installed on a single computer, to be accessed by a group of people at a single location. Other times, they are installed over several servers worldwide, meant to be accessed by millions. These relational enterprise database packages are built and supported by companies such as Oracle, Microsoft, and IBM. The open-source MySQL is also an enterprise database.

As stated earlier, the relational database model does not scale well. The term scale here refers to a database getting larger and larger, being distributed on a larger number of computers connected via a network. Some companies are looking to provide large-scale database solutions by moving away from the relational model to other, more flexible models. For example, Google now offers the App Engine Datastore, which is based on NoSQL. Developers can use the App Engine Datastore to develop applications that access data from anywhere in the world. Amazon.com offers several database services for enterprise use, including Amazon RDS, which is a relational database service, and Amazon DynamoDB, a NoSQL enterprise solution.

2. Making Queries

In general terms, a query is a question, or an inquiry. In a database, a query is a subset of the data that answers a particular question. To show how a query works, and what data is retrieved in response to a query, let's use information from the following database that shows student club membership information.

The following tables in a database are used by the student association of a university which tracks information about club memberships. Queries are made to create a subset of the data if the student association wants to send targeted emails to certain students regarding club events.

Table: Students

Student_ID	Student_Name	Email	On Campus	Major
001	Lee, Peter	p.lee@uni.edu	No	Biology
004	Edwards, Jonathan	j.edwards@uni.edu	Yes	English Literature
007	Johnson, Marilyn	m.johnson@uni.edu	Yes	IT
011	Kim, Joe	j.kim@uni.edu	No	Biology
015	Martinez, Haley	h.martinez@uni.edu	Yes	IT
022	Mfume, John	j.mfume@uni.edu	No	Biology
028	Letty, David	d.letty@uni.edu	Yes	IT
031	Valeriev, Roger	r.valeriev@uni.edu	Yes	Biology
034	Sayer, Harish	h.sayer@uni.edu	Yes	English Literature

Table: Student Clubs

--	--	--	--

Club_ID	Club_Name	Location	Faculty Lead
01	Pre-Med Society	McCormick Hall	Dr. Arntzen
02	Speech & Debate	Norup Hall	Dr. Hendrix
03	Engineering Club	Rice Hall	Dr. Moray
04	Residential Life	Thorton Hall	Dr. Yamasaki
05	Student Health Association	De With Hall	Dr. Dahlmans
06	Math Club	Smith Hall	Dr. Donne

Table: Memberships

Club_ID	Student_ID
01	001
01	022
01	031
02	034
03	007
04	028

The Pre-Med Society is hosting a meeting about special summer internship opportunities for pre-med students. They placed posters around on-campus housing units, but they want to make sure those who live off campus know about this meeting, too.

To make a query, you must first define your criteria. Here, we are interested in all students who: (a) live off campus, and (b) are in the Pre-Med Society. The query applies this criteria to the tables within the database, and displays the subset of all data that meets this criteria:

Student_ID	Student_Name	Email	On Campus	Major
001	Lee, Peter	p.lee@uni.edu	No	Biology
022	Mfume, John	j.mfume@uni.edu	No	Biology

There are two records that meet all of the criteria of the query. Peter Lee and John Mfume are members of the Pre-Med Society and live off campus. The student association can use this list of students who match their criteria to know who to contact about the student club meeting.



SUMMARY

The ability to effectively **manage data** has a huge impact on the overall effectiveness of a business or organization, as decisions are made based on the data available to business leaders. The most important aspect of data management is having access to or being able to retrieve data. **Queries** allow organizations to retrieve specific subsets of data by applying criteria to all of the records in the database tables. Applying effective queries helps organizations to answer specific questions about

the information in their database, and to analyze specific records.

Source: Derived from Chapter 4 of “Information Systems for Business and Beyond” by David T. Bourgeois.
Some sections removed for brevity.

<https://www.saylor.org/site/textbooks/Information%20Systems%20for%20Business%20and%20Beyond/Textbook.html>