

Database Design and SQL

Chapter 1: Introduction

FSDM/CPCM-2023S

Sagara Samarawickrama

Chapter Objectives

- Distinguish between application software and system software and give example for each.
- Distinguish between data and information
- Define database and database management system.
- Describe the steps for the database development process.
- Explain the purpose of conceptual, logical and physical database design.
- Explain the relational database model.

Application Software/System Software

The term '**software**' refers to the set of electronic program instructions or data a computer processor reads in order to perform a task or operation. In contrast, the term '**hardware**' refers to the physical components that you can see and touch, such as the computer hard drive, mouse, and keyboard.

Software can be categorized according to what it is designed to accomplish. There are two main types of software: **systems software** and **application software**.

Application Software/System Software

Systems Software

Systems software includes the programs that are dedicated to managing the computer itself, such as the operating system.

The operating system *manages the computer hardware resources* in addition to applications and data. Without systems software installed in our computers we would have to type the instructions for everything we wanted the computer to do!

Application Software/System Software

Application Software

Application software, or simply *applications*, are often called productivity programs or end-user programs because they enable the user to complete tasks, such as creating documents, spreadsheets, databases and publications, sending emails, designing graphics, and even playing games!

Application software *is specific to the task it is designed for* and can be as simple as a calculator application or as complex as a word processing application



Database Management System(DBMS)

DBMS is a set of software programs that control the organization, storage and retrieval of data in the database.

Technically speaking, DBMS is a software system that uses a standard method of cataloging, retrieving, and running queries on data. The DBMS manages incoming data, organizes it, and provides ways for the data to be modified or extracted by users or other programs.

Common relational database management systems are IBM's DB2, Microsoft SQL Server, Oracle, MySQL, Sybase and others

Relational Database Model

The foundation of most commercial database management systems today is 'Relational database model', which was first introduced in the paper "A relational Model of Data for Large Shared Data Banks" published by Edgar Codd in 1970. Since the publication of this paper the relational model has been developed extensively into common relational database management systems such as Oracle, MySQL, SQL Server and others.

Oracle Database

Oracle database is a relational database management system (RDBMS) from Oracle Corporation. It is one of the most widely-used RDBMS used today

In addition to the relational databases there are other databases such as Hierarchical, Distributed, Object Oriented and NoSQL Databases.

Relational Database Model

All kinds of data are stored in a database. **Data are unorganized raw facts.** When stored in a database, data is stored in rows and columns like a spreadsheet. Each stored item in a database is a raw fact about something.

When data is extracted from a database and processed in some way, it becomes information to the users.

Information is data that has been processed into meaningful results that can be used by users. Information is often the result of combining, comparing, sorting, analyzing or performing calculation of data

Task : Identify a raw facts and information in a customer database

Data vs Information:

Data

- Row Facts
- No context
- Just Numbers and Text

Information

- Data with Context
- Processed data
- Value added to data
 - Analyzed
 - Organized
 - Summarized

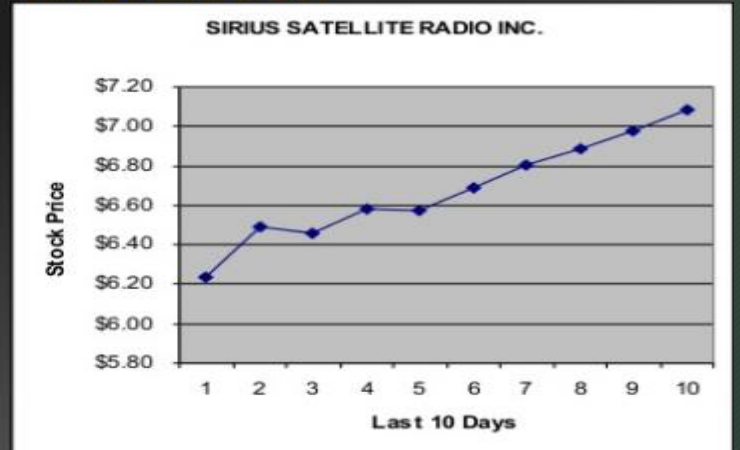
Data vs Information:

- Data : 280717
- Information:
 - 28/07/17 is the date of your final examination.
 - is the average salary for a Business Consultant.
 - is the ZIP code for Metropolis.

Data

- 6.34
- 6.45
- 6.39
- 6.62
- 6.57
- 6.64
- 6.71
- 6.82
- 7.12
- 7.06

Information



Importance of Good Database Design:

The Importance of Good Database Design

Good database design is crucial for a high performance application, just like an aerodynamic body is important to a racecar. If the racecar doesn't have smooth lines, it will produce drag and go slower. The same holds true for databases. If a database doesn't have optimized relationships—normalization—it won't be able to perform as efficiently as possible.

Beyond performance is the issue of maintenance. Your database should be easy to maintain. This includes storing a limited amount (if any) of repetitive data. If you have a lot of repetitive data and one instance of that data undergoes a change (such as a name change), that change has to be made for all occurrences of the data. To eliminate duplication and enhance your ability to maintain the data, you would create a table of possible values and use a key to refer to the value. That way, if the value changes names, the change occurs only once—in the master table. The reference remains the same throughout other tables.

Importance of Good Database Design ...

For example, suppose you are responsible for maintaining a database of students and the classes in which they're enrolled. If thirty-five of these students are in the same class, called “**Database Design**,” this class name would appear thirty-five times in the table. Now, if the instructor decides to change the name of the class to “**Database Programming**”, you must change thirty-five records to reflect the new name of the class. If the database were designed so that class names appeared in one table and just the class ID number was stored with the student record, you would only have to change one record—not thirty-five—in order to update the name change.

The benefits of a well-planned and designed database are numerous, and it stands to reason that the more work you do up front, the less you'll have to do later. A really bad time for a database redesign is after the public launch of the application using it—although it does happen, and the results are costly.

So, before you even start coding your application, spend a lot of time designing your database.

Database Development Process:

The collection of related tables that comprise a database can be quite complex to develop. Because of this, it is important that a set of steps to be followed in the development of a database so users can retrieve the necessary information when needed and in a timely way. The Database Development process includes the following phases

- ❖ Database Planning
- ❖ Requirement Analysis
- ❖ Database Design
- ❖ Database Build
- ❖ Testing and Evaluation
- ❖ Database Implementation
- ❖ Database Maintenance

Database Development Process..

Phase 1 : Database Planning

The database planning phase begins when someone or group **submits a request for the development of a database**. Database planning consists of a set of tasks or activities that decide both the resources required in the database development and the time limits of different activities. During the planning phase, four major activities are performed.

- Review and Approve the database project request
- Prioritize the database project request
- Allocate resources such as money, people and tools
- Arrange a development team to develop the database project.

Database planning should also include the development of the standards that govern how data will be collected, how the format should be specified and what documentation will be needed. (Naming conventions , Data dictionary etc.)

Database Development Process..

Phase 2 : Requirement Analysis

Requirement analysis is the process of gathering the business requirements so they can be used in the development of a data model. This phase involves gathering information about processes, entities and organizational units. After this information is collected, it is used in the database design step. The diagrams produced should show the processes and data that exist as well as the relationships between business processes and data.

During this phase the requirements and the expectations of the users are collected and analyzed. The collected requirements help to understand what the new system will accomplish. The goal of this phase is essentially to understand how the company works, discover what problems and limitations users have, understand what the company wishes to accomplish, and define the scope and boundaries of the project. The scope and boundaries are essential to make sure the database is created exactly as specified.

Database Development Process..

Phase 3 : Database Design

After all the requirements have been gathered for a proposed database, they must be modeled. Database design is the process of producing a detailed data model of the database that will meet end-user requirements. Different types of diagrams are typically produced to illustrate the business processes, rules, entities and organizational units that have been identified. The actual model is frequently called an **entity relational model(ERM)** because it depicts data in terms of the entities and relationships described in the data. An entity relationship model is an abstract conceptual representation of structured data. Basically, data modelling serves as a link between business needs and system requirements.

Designers create an abstract data model that attempt to model real-world objects by creating a conceptual design. Designers must consider end-user views; define entities, attributes and relationships; and identify process and access requirements.

Database Development Process..

Phase 3 : Database Design..

The conceptual design is then translated into logical design , which is DBMS-dependent. Database design includes three design steps:

1. Conceptual Design

2. Logical Design

3. Physical Design

This significance of this approach is that it allows the three perspectives to be relatively **independent of each other**. Storage technology can change without affecting either the conceptual or logical model.

Database Development Process..

Phase 3 : Database Design..

Conceptual Design:

The conceptual design is the process of constructing a high-level graphical representation of the proposed database showing entities and relationships in the form of an ERD from the information collected in the requirement analysis phase. It does not contain details of the entities and relationships in detail.

The conceptual model is a representation of an organization's data, organized in terms of entities and relationships and is **independent of any particular DBMS**.

Database Development Process..

Phase 3 : Database Design..

Logical Design:

The logical data model is a continuation of the conceptual data model and represent the entities, attributes and relationships in detail. The logical data model is the model from which the physical layer design is created.

Logical design consists of expanding the entity relationship diagram from the conceptual model into a more detailed pictorial representation of the real world problem in terms of entities, attributes and relations between the entities.

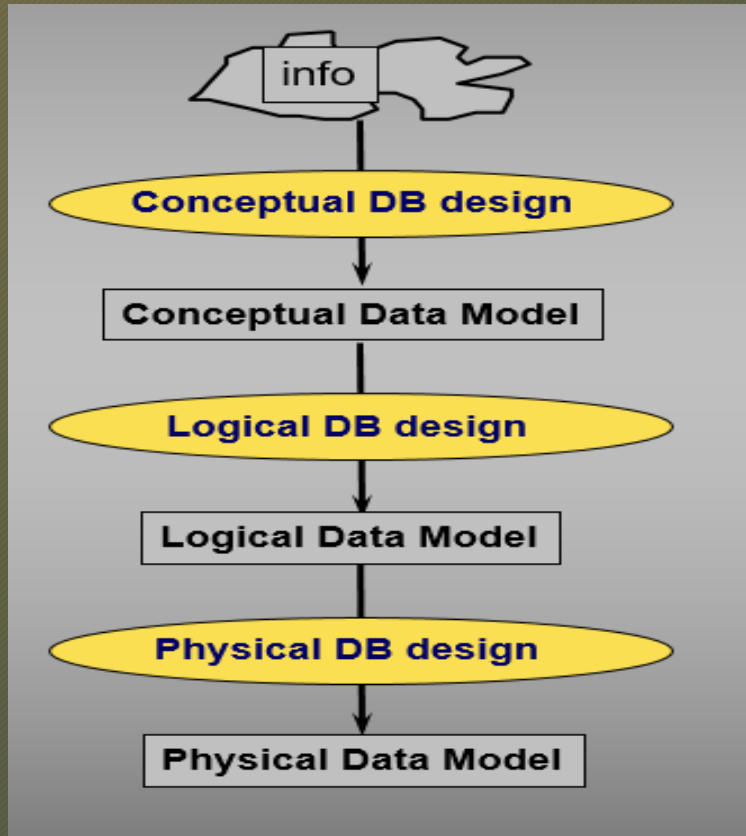
Entities : people, places, things or concepts about which the data must be recorded such as customers or employees.

Attributes : Attributes are properties or characteristics of entities such as customer names , dates on which orders were placed.

Database Development Process..

Phase 3 : Database Design..

Relationships : Relationships represent association between or among entities.



Normalization : is an important part of a database and is used to eliminate redundancy and other anomalies in the database

Database Development Process..

Phase 3 : Database Design..

Physical Design:

Physical design is the transformation of the logical data model representing entities, attributes and relationships into a physical model representing database tables, columns, and relationships. The result is a physical data model that represents database tables, columns and relationships that are ready for implementation.

Once the data model for the database is completed ,implementation follows

DBMS Selection: In this phase , an appropriate DBMS is selected to support the information system. A number of factors are involved in DBMS Selection, including technical and economic factors. Technical factors include type of DBMS, storage structure and access method, type of query language and user/program interface

Database Development Process..

Phase 4 : Database Build

After the design phase and selection of a suitable DBMS , the database systems built.

The purpose of the database build phase is to construct and install the database according to the plan and design as described earlier.

In this phase the database management system software is installed , the database is created with all the components

Database Development Process..

Phase 5 & 6: Database Implementation & Testing/Evaluation

Database Implementation:

The purpose of the implementation phase is to construct and install the database according to the plan and design as described in previous phases. In this phase, the DBMS software is installed, the database is produced, and the tables populated with data. This phase also requires that the database performance is evaluated, security standards are setup, back up and recovery procedures are put in place, and data integrity is enforced.

Testing and Evaluation:

This phase requires that the database is tested again for performance. Although it is tested during implementation phase, in this phase it is re-tested and fine-tuned. Testing also requires that the administrator test integrity, security and multiuser load.

Database Development Process..

Phase 7: Database Maintenance

Database Maintenance:

Once the database system is implemented, the maintenance phase of the database system begins. Database maintenance is the process of monitoring and maintaining the database system. Maintenance include activities such as adding new columns to the table, changing the size of the existing columns, adding new tables, adding new constraints and so on. Some queries or transactions may be rewritten for better performance.

Introduction to SQL

SQL pronounced "S-Q=L" stands for Structured Query Language is the industry-standard language of relational database management systems (RDBMS) for manipulating and querying data in relational databases.

SQL , originally designed by IBM in the mid-1970s, came into widespread use in the early 1980s. In 1986 , SQL became an industry standard of American National Standards Institute(ANSI), and the International Organization for Standardization (ISO) in 1987.Despite the existence of ANSI/ISO standard, database vendors have included proprietary extensions that have resulted in different versions of SQL Language.

Introduction to SQL

SQL is the language of the database management system. Some of the functions performed by SQL include

- ✓ Create and Delete Tables in a database
- ✓ Run queries against a database
- ✓ Retrieve data from a database
- ✓ Insert Rows of data to table
- ✓ Update existing data
- ✓ Delete rows from a table
- ✓ Create Views

Introduction to SQL

SQL Commands are mainly categorized into four categories

DDL (Data Definition Language)

DML (Data Manipulation Language)

DCL (Data Control Language)

TCL (Transaction Control Language)

DDL (Data Definition Language)

DDL Commands used to create and modify the structure of database objects.

Introduction to SQL

Command	
CREATE	Create database and objects (table, view, index, and others)
DROP	Delete objects from the database
ALTER	Alter the structure of the database
TRUNCATE	Remove all rows (records) from a table
COMMENT	Add comments to the data dictionary
RENAME	Rename an object existing in the database

DML (Data Manipulation Language)

DML commands that deal with the manipulation of data in a database

Command	
SELECT	Returns a dataset from a database table based on selection criteria
INSERT	Inserts (adds) a row into a database table
UPDATE	Updates existing data in one or more rows in a database table
DELETE	Removes one or more rows from a database table

Introduction to SQL

DCL (Data Control Language)- DCL commands that deal with the rights permissions and other controls of the database systems

Command	
GRANT	Gives user's access privileges to database
REVOKE	Withdraw user's access privileges given by using the GRANT command

Transaction Control Systems (TCL) -TCL commands used to manage changes made to the data in a table by DML statements.

Command	
COMMIT	Commits a transaction
ROLLBACK	Rollbacks or reverses a transaction when errors occur
SAVEPOINT	Sets a SAVEPOINT within a transaction
SET TRANSACTION	Specifies characteristics for a transaction

Introduction to SQL

Creating Schema

A Schema must be created before the creation of tables and other objects. A Schema is an object that serves as a container for database objects such as tables , views , indexes , stored procedures and other object types.

```
CREATE SCHEMA myschema;
```

The **SET SCHEMA** command changes the default schema for all subsequent SQL Commands. To Change the default schema enter the keywords **SET SCHEMA** followed by the name of the Schema

```
SET SCHEMA myschema;
```

Conclusion

