Introduction to Databases

Oscar Rodríguez de Rivera, PhD

What is a Database?

A database is a collection of organised data, information and record.

Purpose of a Database

Database is information that a person needs in his personal, business, social, etc. and the power and purpose of information is not only in collecting and finding them but more importantly using them.

Kinds of Databases

Structured database

It is also called the structured data in which a record or file of information arranged in uniform format. These database are usually storage of information with similar entries such as a list of person born in a country, medical data base of patients, inventory, etc.

Free-form database

It is a loose collection of information, such as those you will find no the WWW. A collection of your documents in the computer made from several programs can be considered as free-form database.

Types of Databases

Operational database

It is a dynamic database that is used by any organisation in its day-to-day operation. They are used to collect data, maintain, modify and manage data.

Analytical database

It is a static database, where data is rarely modified. This database is often used to store and track historical data to make long term projections and analysis.

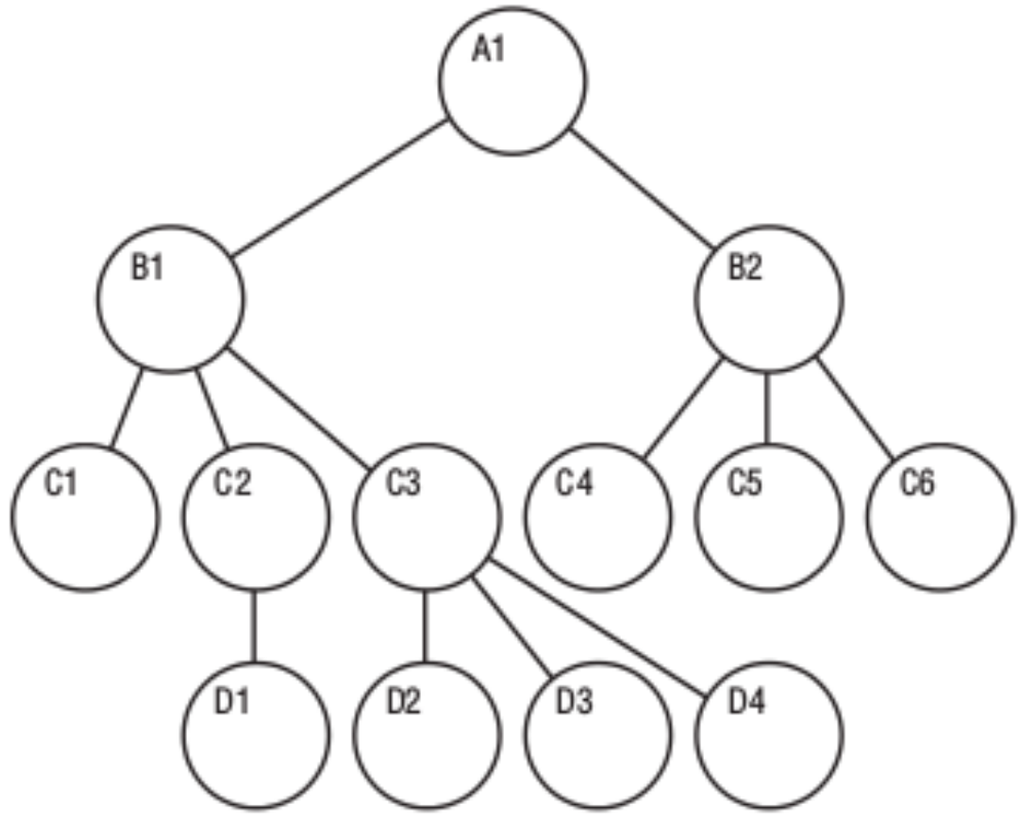
Structured Database Models

Hierarchical Model

This model can be visualised as a parent-child relationship wherein a child my only have one parent but parent can have several children.

Another way of looking at this model is by visualising an inverted tree. The single table acts as the root of the inverted tree and the other tables act as branches.

To access the data from one of the tables, one has to pass through the root table.

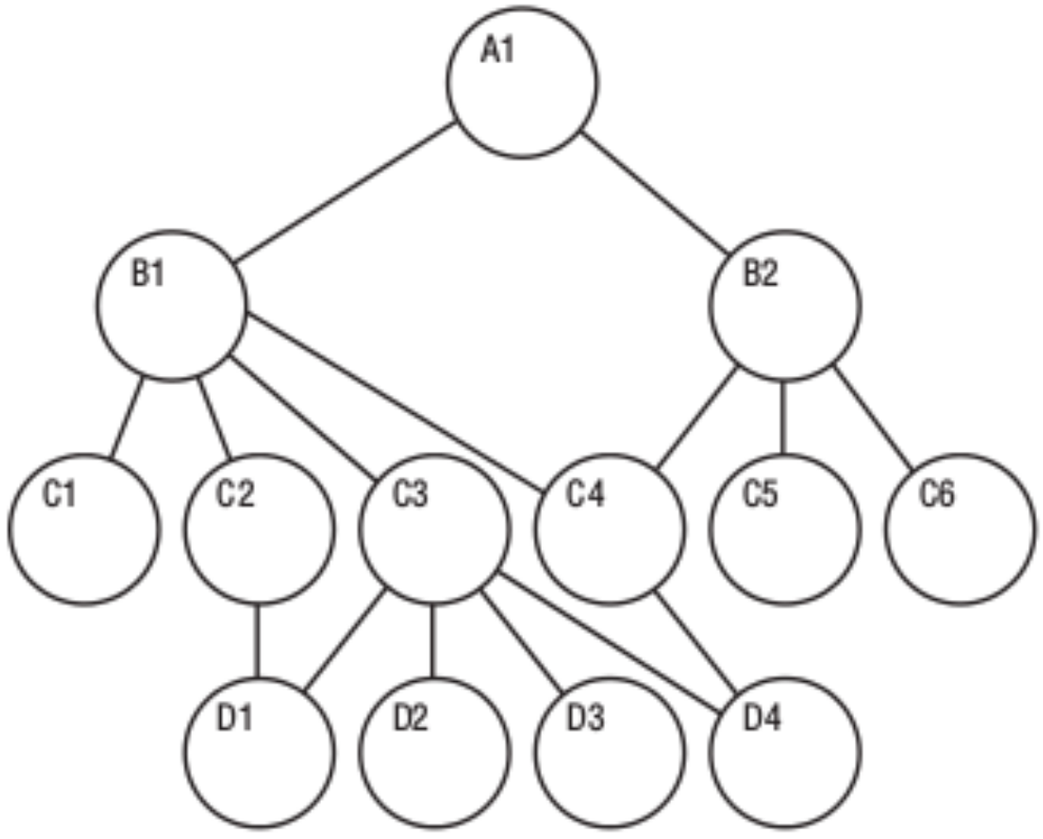


Network Model

This model was developed to address in part the problem of the hierarchical model. As with hierarchical model, it could be visualised as an inverted tree; however, this models allows many inverted trees sharing branches but are still part of the same database structure.

As with the parent-child model, the child is allowed to have multiple parents. Unlike in the hierarchical model access data begins from the root table.

In the network model, it allows access of data from any table. This model was eventually replaced by the relational debase model.



Relational Model

Was developed to solve the problems of the earlier database models.

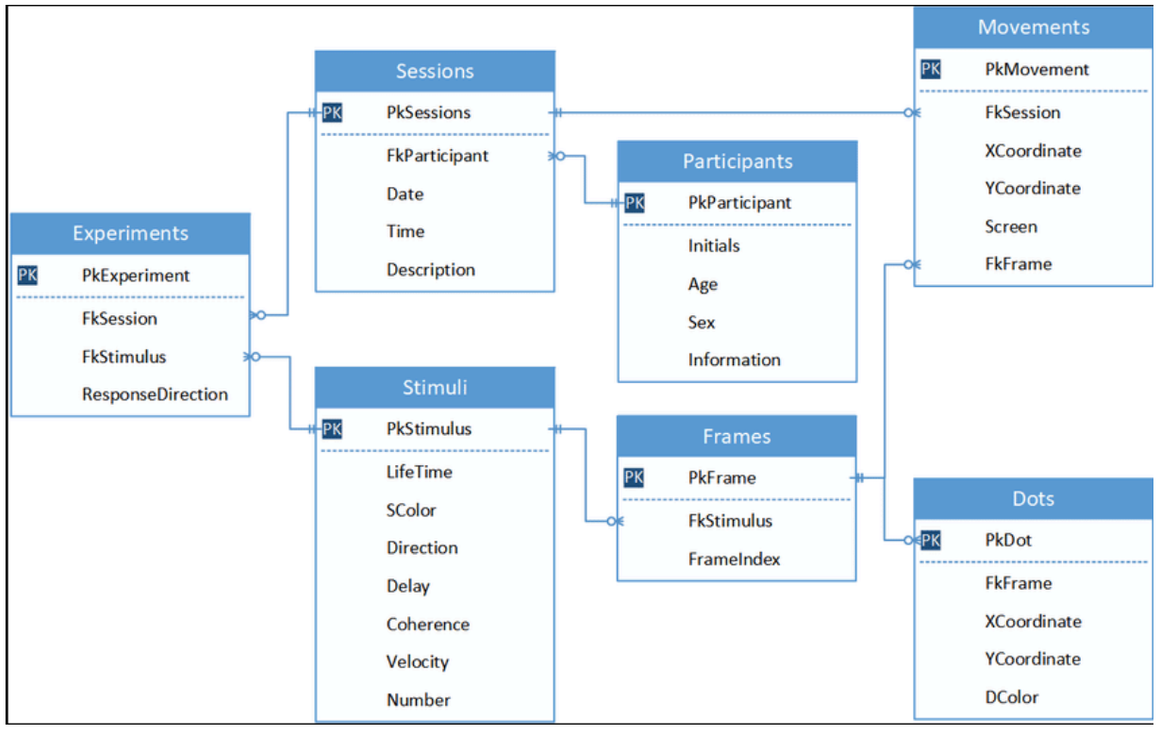
The terms relations, attributes and domains are used frequently in relational debase models.

A relation is a table with columns & rows.

Attributes are the names given to each column of the relation.

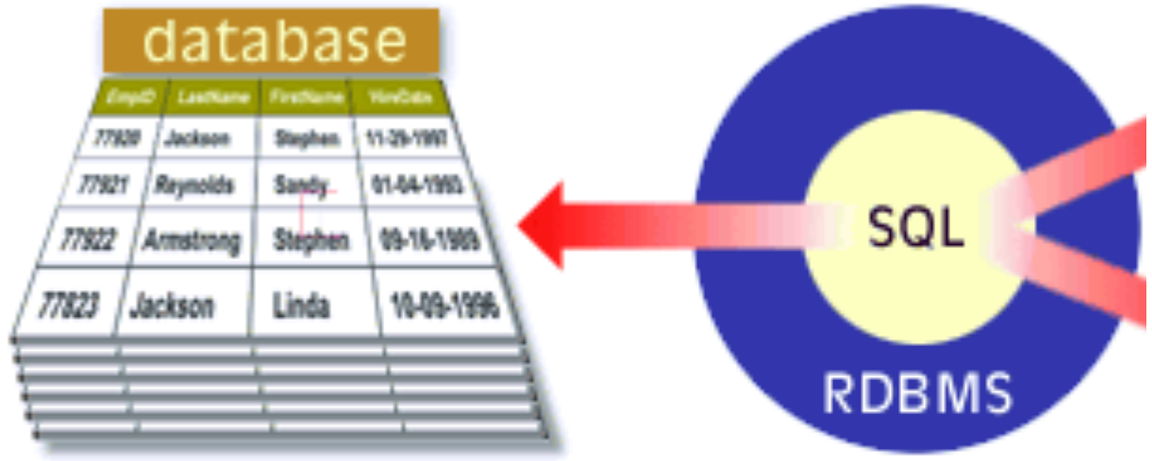
Domain is the value with which the attribute will take.

A relational database is essentially a collection of tables, and if a database is a collection of tables, these tables are simply called relations, hence, the name.



Relational Database Management System

RDBMS is designed to create, maintain, manipulate, modify and delete information in a relational database



Modern database use the relational database model and many of today's software caters to this type of structured database.

Things you can do with RDBMS

- Create a database
- Information storage
- Information retrieval
- Information management
- Information analysis
- Print and share information

Relational Database Terminologies

Data

It is a number or value found and stored in the database. Data is static because it remains the same until it is modified by a process.

Information

It is data that has been processed thereby making it. Relevant and meaningful to the person viewing it. Information is dynamic because it changes relative to the data stored in the database and it could be processed in many ways.

Null

It is used to represent a value that is unknown or missing. A null value is neither a zero nor a blank.

Table

It is the main structure in the relational database. It is composed of attributes (fields) and domain (records). A table almost always represents a subject that can be an object (person, place, or thing) or an event.

File

It is an organised collection of data about an entity. As an example, for a bookstore, a file called “Branch” can contain all the data about a particular bookstore branch.

Record

It refers to a specific person, place, thing, or event. Record is also known as the “tuple” in the relational database terminology. It pertains to structure in the dabs table representing a unique instance of a subject.

Field

It is the smallest structure of a data from a larger database structure in a relational database. A field can store data in a database and represent a character of the subject to which database table it resides.

View

It is also known as a virtual table. It is called a virtual table since does not hold data on its own; rather it gets data from the table which it is based. And since it comes from other table it is composed of several fields coming from one or more data.

Keys

These are fields that serve specific purposes within a table. There are two types of keys, the primary key and the foreign key.

The Primary key is a field that uniquely identifies a record in the table.

The Foreign key is a special field that establishes a relationship between two tables

Relationship

They exist when two or more tables have connection or association.

Relationships

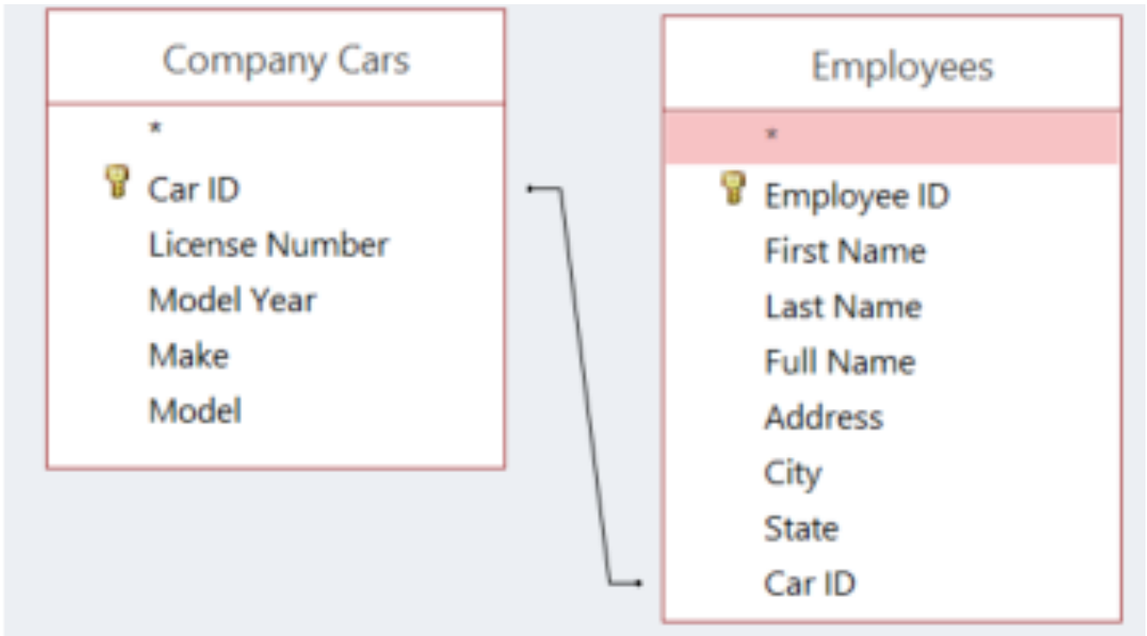
- Primary key
- Foreign key
- Linking table is a table that establishes a connection between two or more tables.

Type of Relationships

When two or more tables are related, there exist between them a specific type of relationship and there are three types of possible relationships.

One-to-One relationships

This relationship exist between tables tables when only one record of the first table is related to only one record to a second table, and only one record of the second table is related to only one record to the first table.



One-to-Many relationships

This relationship exist between tables when one record of the first table can be related to one or more records to a second table, but only one record from the second table can be related to a single record in the first table.

This relationship is the most common relationship that exist between tables and helps to reduce or eliminate duplicate and redundant data.

Customers

Customer ID	CustFirst Name	CustLast Name	*<< other fields >>*
9001	Paul	Litwin	.....
9002	Alison	Balter	.....
9003	Andy	Baron	.....
9004	Chris	Kunicki	.....
9005	Mary	Chipman	.....

Customer Rentals

Customer ID	Video ID	Checkout Date
9002	80115	09/26/01
9001	64558	09/28/01
9003	10202	09/28/01
9003	11354	09/28/01
9003	78422	10/02/01
9005	30556	09/26/01
9004	20655	10/05/01

Many-to-Many relationships

This relationships exists between tables when one record of the first table can be related to one or more records to a second table and one record from the second table can be related to one or more records to the first table.

The connection between the two tables will be difficult to establish and will result to redundant data in one of the tables.

Classes		
Class Name	Credit Hours	Class ID
Introduction to Jazz	4	101
Introduction to Guitar	5	102
Advanced Guitar	4	103
Jazz Band	4	104

Students		
Student ID	Student Name	Date of Birth
1001	Wayne Gentry	3/5/2005
1002	Cathy Baker	10/2/2005
1003	Zayan Dean	12/3/2004
1004	Isa Mcleod	4/18/2005
1005	Jud Wilks	6/4/2005
1006	Sally Driscoll	10/3/2004

Enrollments			
Class ID	Student ID	Class Grade	Enrollment ID
101	1001	B+	100001
101	1003	A-	100002
101	1004	B	100003
101	1006	A+	100004
102	1001	C+	100005
102	1004	A+	100006
102	1005	C	100003
103	1002	B	100004
103	1006	B-	100005
104	1002	A	100006
104	1005	A-	100006
104	1006	B+	100006

Designing a database

- Define the purpose of your database. Consider the questions or queries you may want to answer about the stored data.
- Determine the tables that you need in the database.
- Determine the fields that you need in the database.
- Identify unique fields values that you will allow to connect information stored in a separate table.
- Determine relationships between tables. A relationship works by matching data in the key fields which is usually a field with the same name in both tables.
- Test the design by entering the sample data. Check that you can run a query on the database and get the information you want.

Characteristics of a Well-Designed Database

- Modify data is easy. Changes to the value of one field within the table should not affect the values of the fields in the table.
- Retrieving information is easy. Extracting desired information from tables with well defined relationships should make accessing and retrieving data a lot faster.
- Developing and building user application is easy. Data manipulation would be the main focus of programming and not solving the problems associated with a poorly designed database.
- Maintaining the structure is easy. Changes made to any table or columns should not affect other tables or columns.
- Adding and deleting data is easy.

Database Designs to Avoid

Spreadsheet Design

A spreadsheet is very powerful if used properly; however, it is designed for a purpose and always will find a purpose within any business or organisation.

Spreadsheets do not make good relational database, because if you really need to acquire and collect, store and maintain, view and analyse, print and share data then you need a tool that will truly suit in designing a real database.

Spreadsheet View

Secondly, you should also get away from the “spreadsheet view” mentality. Spreadsheets may seem to offer a good means to view data; however, a database program may not be able to reproduce a report of a spreadsheet layout.

A database presentation may not be the same as that of a spreadsheet but it could be just as clear.