Database Schema

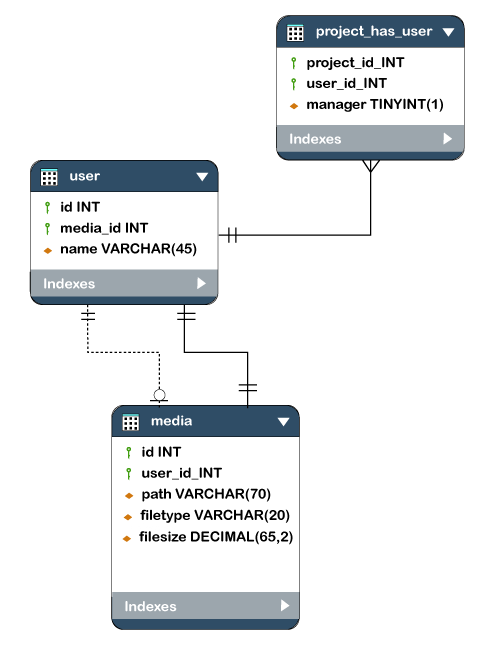
**A database schema is a structure that represents the logical storage of the data in a database**. It represents the organization of data and provides information about the relationships between the tables in a given database. In this topic, we will understand more about database schema and its types. Before understanding database schema, lets first understand what a Database is.

What is Database?

A database is a place to store information. It can store the simplest data, such as a list of people as well as the most complex data. The database stores the information in a well-structured format.

What is Database Schema?

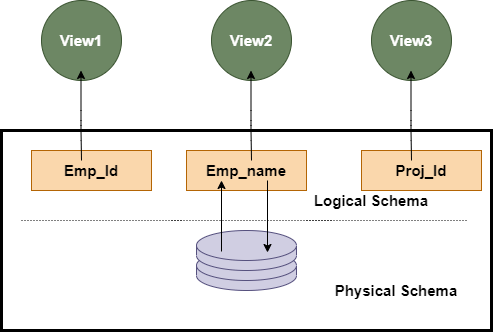
* A database schema is the logical representation of a database, which shows how the data is stored logically in the entire database. It contains list of attributes and instruction that informs the database engine that how the data is organized and how the elements are related to each other.
* A database schema contains schema objects that may include **tables, fields, packages, views, relationships, primary key, foreign key,**
* In actual, the data is physically stored in files that may be in unstructured form, but to retrieve it and use it, we need to put it in a structured form. To do this, a database schema is used. It provides knowledge about how the data is organized in a database and how it is associated with other data.
* **The schema does not physically contain the data itself; instead, it gives information about the shape of data and how it can be related to other tables or models.**
* A database schema object includes the following:
  + Consistent formatting for all data entries.
  + Database objects and unique keys for all data entries.
  + Tables with multiple columns, and each column contains its name and datatype.
* The complexity & the size of the schema vary as per the size of the project. It helps developers to easily manage and structure the database before coding it.
* The given diagram is an example of a database schema. It contains three tables, their data types. This also represents the relationships between the tables and primary keys as well as foreign keys.



Types of Database Schema

The database schema is divided into three types, which are:

1. **Logical Schema**
2. **Physical Schema**
3. **View Schema**



1. Physical Database Schema

A physical database schema specifies how the data is stored physically on a storage system or disk storage in the form of Files and Indices. **Designing a database at the physical level is called a physical schema.**

2. Logical Database Schema

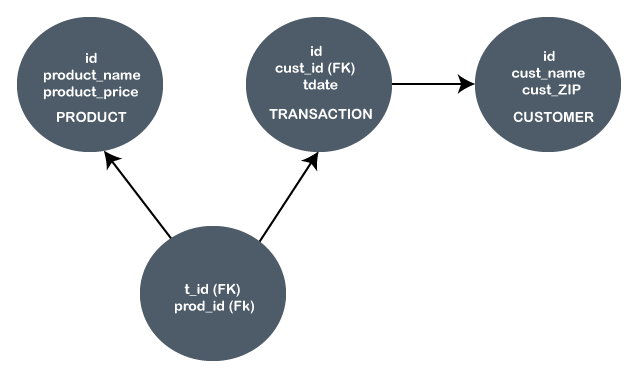
The Logical database schema specifies all the logical constraints that need to be applied to the stored data. It defines the views, integrity constraints, and table. Here the term **integrity constraints** define the set of rules that are used by DBMS (Database Management System) to maintain the quality for insertion & update the data. The logical schema represents how the data is stored in the form of tables and how the attributes of a table are linked together.

At this level, programmers and administrators work, and the implementation of the **data structure** is hidden at this level.

Various tools are used to create a logical database schema, and these tools demonstrate the relationships between the component of your data; this process is called **ER modelling**.

The ER modelling stands for entity-relationship modelling, which specifies the relationships between different entities.

We can understand it with an example of a basic commerce application. Below is the schema diagram, the simple ER model representing the logical flow of transaction in a commerce application.



In the given example, the Ids are given in each circle, and these Ids are primary key & foreign keys.

The **primary key is** used to uniquely identify the entry in a document or record. The Ids of the upper three circles are the primary keys.

The **Foreign key** is used as the primary key for other tables. The FK represent the foreign key in the diagram. It relates one table to another table.

3. View Schema

The view level design of a database is known as **view schema**. This schema generally describes the end-user interaction with the database systems.

Difference between the Physical and Logical Database Schema

|  |  |
| --- | --- |
| **Physical database schema** | **Logical Database schema** |
| It does not include the attributes. | It includes the attributes. |
| It contains both primary & secondary Keys. | It also contains both primary & secondary keys. |
| It contains the table name. | It contains the names of the tables. |
| It contains the column names and their data types. | It does not contain any column name or datatype. |

Database Instance or Database Schema is the same?

The terms database schema and database instances are related to each other & sometimes confusing to be used as the same thing. But both are different from each other.

**Database Schema** is a representation of a planned database and does not actually contain the data.

On the other hand, a **database instance** is a type of snapshot of an actual database as it existed at an instance of time. Hence it varies or can be changed as per the time. In contrast, the database schema is static and very complex to change the structure of a database.

Both instances and schemas are related to and impact each other through the DBMS. DBMS ensures that every database instance complies with the constraints imposed by the database designers in the database schema.

Creating Schema

To create a schema, "CREATE SCHEMA" Statements is used in each type of database. But each DBMS has a different meaning for this. Below we are explaining creating schema in different database systems:

**1. MySQL**

In MySQL, the "**CREATE SCHEMA**" statement creates the database. It is because, in MySQL, the CREATE SCHEMA statement is similar to CREATE DATABASE statement, and schema is a synonym for the database.

**2. Oracle Database**

In Oracle Database, each schema is already present with each database user. Hence CREATE SCHEMA does not actually create a schema; rather, it helps to show the schema with tables and views and allows to access those objects without requiring multiple SQL statements for multiple transactions. The "CREATE USER" statement is used to create a schema in Oracle.

**3. SQL Server**

In the SQL server, the "CREATE SCHEMA" statement creates a new schema with the name provided by the user.

Database Schema Designs

A schema design is the first step in building a foundation in data management. Ineffective schema designs are difficult to manage and consume more memory and other resources. It logically depends on the business requirements. It is required to choose the correct database schema design to make ease in the project lifecycle. The list of some popular database schema designs is given below:

* **Flat Model**
* **Hierarchical Model**
* **Network Model**
* **Relational Model**
* **Star Schema**
* **Snowflake Schema**

Flat Model

A flat model schema is a type of 2-D array in which each column contains the same type of data, and elements within a row are related to each other. It can be understood as a single spreadsheet or a database table with no relations. This schema design is most suitable for small applications that don't contain complex data.

Hierarchical Model

The Hierarchical model design contains a tree-like structure. The tree structure contains the root node of data and its child nodes. Between each child node and parent node, there is a one-to-many relationship. Such type of database schemas is presented by XML or JSON files, as these files can contain the entities with their sub-entities.

The hierarchical schema models are best suitable for storing the nested data, such as representing **Hominoid classification.**

Network Model

The network model design is similar to hierarchical design as it represents a series of nodes and vertices. The main difference between the network model and the hierarchical model is that the network model allows a many-to-many relationship. In contrast, the hierarchical model only allows a one-to-many relationship.

The network model design is best suitable for applications that require spatial calculations. It is also great for representing workflows and mainly for cases with multiple paths to the same result.

Relational Model

The relational models are used for the relational database, which stores data as relations of the table. There are relational operators used to operate on data to manipulate and calculate different values from it.

Star Schema

The star schema is a different way of schema design to organize the data. It is best suitable for storing and analysing a huge amount of data, and it works on "Facts" and "Dimensions". Here **the fact** is the numerical data point that runs business processes, and **Dimension** is a description of fact. With Star Schema, we can structure the data of RDBMS.

Snowflake Schema

The snowflake schema is an adaption of a star schema. There is a main "Fact" table in the star schema that contains the main data points and reference to its dimension tables. But in snowflake, dimension tables can have their own dimension tables.