

Introducción a las Bases de Datos

Generalidades

La definición de bases de datos es “aquel conjunto de datos almacenados y estructurados según sus características o tipología para ser utilizados o consultados posteriormente”.

Otra definición según Oracle, una base de datos es una recopilación organizada de información o datos estructurados, que normalmente se almacena de forma electrónica en un sistema informático. Normalmente, una base de datos está controlada por un sistema de gestión de bases de datos (DBMS). En conjunto, los datos y el DBMS, junto con las aplicaciones asociadas a ellos, reciben el nombre de sistema de bases de datos, abreviado normalmente a simplemente base de datos.

Los datos de los tipos más comunes de bases de datos en funcionamiento actualmente se suelen utilizar como estructuras de filas y columnas en una serie de tablas para aumentar la eficacia del procesamiento y la consulta de datos. Así, se puede acceder, gestionar, modificar, actualizar, controlar y organizar fácilmente los datos. La mayoría de las bases de datos utilizan un lenguaje de consulta estructurada (SQL) para escribir y consultar datos.

Tipos de Bases de Datos

Existen muchos tipos distintos de bases de datos. Lo más habitual es clasificarlas según el modelo, su contenido o la variabilidad de los datos que incluyen. Pasamos a ver los tipos más frecuentes de bases de datos.

Según el modelo

La primera clasificación de bases de datos es según el modelo empleado. En este caso podemos encontrar bases de datos relacionales, distribuidas, orientadas a objetos, documentales, deductivas, transaccionales y más.

- **Relacionales**

Las bases de datos relacionales son unas de las más frecuentes por su flexibilidad y facilidad de uso. En este modelo no importa el lugar o la forma en la que estén almacenados los datos. Por el contrario, se accede a la información mediante consultas que permiten acceder de forma ágil y flexible a los datos. Suelen emplear el lenguaje SQL.

- **Distribuida**

Las bases de datos distribuidas consisten en bases de datos establecidas en diferentes lugares y conectadas por una red. Se emplean en organismos descentralizados que necesitan unir bases de datos de diferentes lugares (localidades, universidades, etc). Existen bases distribuidas de forma homogénea (usan el mismo SGDB) o heterogénea (emplean sistemas multibase).

- **NoSQL**

Son bases de datos que no usan el lenguaje SQL, o que lo usan solo como apoyo, pero no como consulta. Entre los lenguajes más usados por las bases de datos no relacionales están CQL (Contextual Query Language, JSON (JavaScript Object Notation) y GQL (Graph Query Language).

- **Orientadas a objetos**

Las bases de datos orientadas a objetos están enfocadas al almacenamiento de objetos completos, incluyendo su estado y comportamiento. Este tipo de bases de datos se basan en conceptos básicos del tratamiento informatizado de objetos, como la herencia, encapsulación o polimorfismo.

- **Multidimensionales**

Las bases de datos multidimensionales son similares a las relacionales, pero se suelen usar para el desarrollo de aplicaciones muy específicas. La principal diferencia entre ambas es que en las multidimensionales los atributos de las tablas pueden representar tanto dimensiones como métricas.

- **Documentales**

Las bases de datos documentales están pensadas para indexar textos completos y, por tanto, han de estar preparadas para almacenar una gran cantidad de información.

- **Deductivas**

Las bases de datos deductivas o lógicas se asientan sobre la lógica matemática y son capaces de hacer deducciones basadas en hechos o reglas establecidas, Gracias a ello permiten establecer relaciones indirectas entre diferentes tipos de datos y solventar algunas de las limitaciones de las bases de datos relacionales.

- **Transaccionales**

Las bases de datos transaccionales se emplean para enviar y recibir datos a gran velocidad. En este tipo de bases de datos no se tiene en cuenta la redundancia o duplicidad de los datos ya que su función es exclusivamente el envío y recepción de información a la mayor velocidad posible.

- **Jerárquicas**

Las bases de datos jerárquicas se organizan en base a un nudo principal de información, o nodo padre (raíz), del cual surgen diferentes hijos (hojas). Es un sistema que organiza los datos de forma jerárquica, en forma de árbol invertido.

- **Red**

Las bases de datos de red son similares a las jerárquicas, con la salvedad de que pueden existir varios nodos padres. Esto supuso un importante avance respecto a las bases jerárquicas, sobre todo en lo relativo a la redundancia de datos.

Según el contenido

Otro método para clasificar las bases de datos es según la forma en que presentan el contenido. En este caso nos podemos encontrar con dos tipos, las bibliográficas y las de texto completo.

- **Bibliográficas**

Las bases de datos bibliográficas muestran información relevante sobre una determinada publicación: título, autor, año de publicación, editorial, etc. También pueden presentar un pequeño extracto del texto, pero nunca la publicación original al completo, ya que sino estaríamos hablando de la tipología que sigue a continuación.

- **Texto completo**

Las bases de datos de texto completo guardan y muestran todo el contenido de las publicaciones. Por ejemplo, todo el contenido de los números publicados de una determinada revista.

Según la variabilidad

Otra de las variables que se emplean para la clasificación de bases de datos es la variabilidad de la información que presentan. Así, podemos distinguir entre databases estáticas o dinámicas.

- **Estáticas**

Las bases de datos estáticas almacenan datos fijos que no se modifican con el tiempo. Normalmente se trata de datos históricos que pueden ser estudiados para ver su evolución en el tiempo y realizan proyecciones o tomar decisiones en base a dicha evolución.

- **Dinámicas**

Por su parte, las bases de datos dinámicas almacenan información que se modifica con el tiempo. Los datos se editan y actualizan a medida que van cambiando. Por ejemplo, una base de datos con los precios de un comercio.

Características

Las principales características de una base de datos son las siguientes:

- Los datos almacenados tienen independencia física y lógica.
- Garantizan la integridad de los datos.
- Son sistemas de almacenamiento que contribuyen a reducir la redundancia al mínimo.
- Es habitual que puedan acceder a ellas múltiples usuarios.
- Permiten realizar consultas complejas de diferentes tipos de datos.
- El acceso a los datos es seguro
- Se pueden hacer copias de seguridad y recuperar los datos
- Se accede a la información a través de un lenguaje concreto.

Sistemas Gestores de Bases de Datos

Normalmente, una base de datos requiere un programa de software de bases de datos completo, conocido como sistema de gestión de bases de datos (DBMS). Un DBMS sirve como interfaz entre la base de datos y sus programas o usuarios finales, lo que permite a los usuarios recuperar, actualizar y gestionar cómo se organiza y se optimiza la información. Un DBMS también facilita la supervisión y el control de las bases de datos, lo que permite una variedad de operaciones administrativas como la supervisión del rendimiento, el ajuste, la copia de seguridad y la recuperación.

Algunos ejemplos de software de bases de datos o DBMS populares incluyen MySQL, Microsoft Access, Microsoft SQL Server, FileMaker Pro, Oracle Database y dBASE.

Estructuras en una Base de Datos



Tablas

Dentro de cada base de datos se crearán, primeramente, las tablas. Las tablas son otras estructuras -siempre dentro de bases de datos, nunca fuera de ellas- que guardan un subconjunto del todo. En el ejemplo de la escuela, podríamos tener una tabla con los datos de los alumnos, otra tabla con los datos de los profesores, otra con los planes de estudio y así hasta tener todos los datos de control que queramos guardar en la base de datos. Si pensáramos en términos de archivos de hojas de cálculo, cada hoja de cálculo sería una tabla (que luego veremos que se conforman de registros y columnas) y la carpeta o subdirectorío donde se guardan todas las hojas de cálculo sería la base de datos.

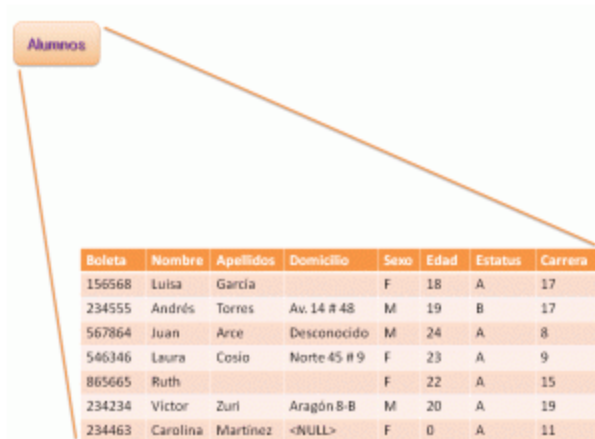


Diagrama de una tabla de alumnos. El título "Alumnos" está en la parte superior. Debajo, se listan los atributos y sus tipos de datos en un formato de lista con viñetas.

Alumnos	
•Boleta	(Texto)
•Nombre	(Texto)
•Apellidos	(Texto)
•Domicilio	(Texto)
•Sexo	(Clave)
•Edad	(Entero)
•Estatus	(A/B)
•Carrera	(clave)

Registros y columnas

Dentro de cada tabla todos los registros tienen el mismo formato, es decir, tienen los mismos campos. Siguiendo con el ejemplo de control escolar, y pensando en la tabla de alumnos, podríamos definir que cada alumno tendrá los mismos 10 campos o atributos (número de boleta, nombre completo, fecha de nacimiento, escuela de origen, promedio actual, carrera que cursa, correo electrónico, teléfono de contacto, domicilio y nombre del padre o tutor). Todos y cada uno de los alumnos contará con espacio para guardar esos 5 atributos, ni un dato más ni un dato menos, todos tendrán la misma estructura. Utilizando el ejemplo de la hoja de cálculo como simil de una tabla, la hoja de cálculo tendrá las mismas 10 columnas para cada alumno que guardemos en dicha hoja de cálculo



Boleta	Nombre	Apellidos	Domicilio	Sexo	Edad	Estatus	Carrera
156568	Luisa	García		F	18	A	17
234555	Andrés	Torres	Av. 14 # 48	M	19	B	17
567864	Juan	Arce	Desconocido	M	24	A	8
546346	Laura	Cosío	Norte 45 # 9	F	23	A	9
865665	Ruth			F	22	A	15
234234	Victor	Zuri	Aragón 8-B	M	20	A	19
234463	Carolina	Martínez	<NULL>	F	0	A	11

Tipos de Datos

¿Qué es un dato?

Un dato nos permite describir un objeto. Dicho objeto podemos llamarlo entidad, por ejemplo una casa en la que viven personas. La casa es la entidad y la cantidad de personas que viven en la casa son un dato, que en este caso es numérico.

Hay diferentes tipos de datos que se pueden tener en una base de datos: caracteres, numéricos, imágenes, fechas, monedas, texto, bit, decimales y varchar. Hay bases de datos más o menos adecuadas según el tipo de dato, y hay un proceso para convertir un dato en información, pues un solo dato por sí mismo no representa nada si no lo vemos en contraste con otros. Así se identifican comportamientos.

Relaciones en Base de Datos

Para saber qué son las relaciones en las bases de datos es necesario hablar del tipo de base de datos relacionales. En estas bases de datos, la información se almacena en diferentes tablas, distribuida en filas y columnas.

La relación de una base de datos es el vínculo que se establece entre distintos elementos de las tablas que la conforman. En este tipo de relaciones es fundamental el uso de los campos de llave primaria (primary key) que son los que se relacionan con otros registros de otras tablas.

Es importante destacar que, a la hora de definir las relaciones entre los campos de distintas tablas en una base de datos, los nombres de los mismos no tienen por qué ser iguales. Sin

embargo, sí es necesario a la hora de establecer estas relaciones, que el tipo de datos de los campos enlazados sea el mismo.

Para qué sirven

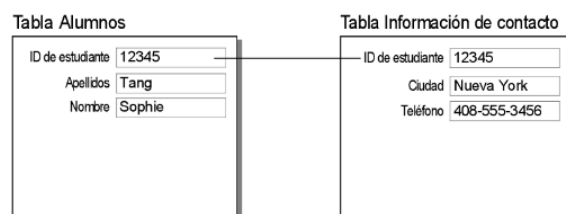
Las relaciones en las bases de datos son claves para establecer concordancias en las asignaciones y garantizar la integridad referencial de la información (que los datos no se modifiquen o varíen durante el proceso).

Gracias a las relaciones se mantiene una lógica y consistencia entre todos los datos que almacena. Además, las relaciones evitan que se dupliquen los registros dentro de una base de datos.

Relación uno a uno (1:1)

Se produce cuando la relación se realiza solo entre un registro de una tabla con un registro de otra. Es una de las relaciones más utilizadas, ya que permiten una relación de tipo exclusivo. En esta relación, los campos establecidos como primary key de ambas tablas están enlazados.

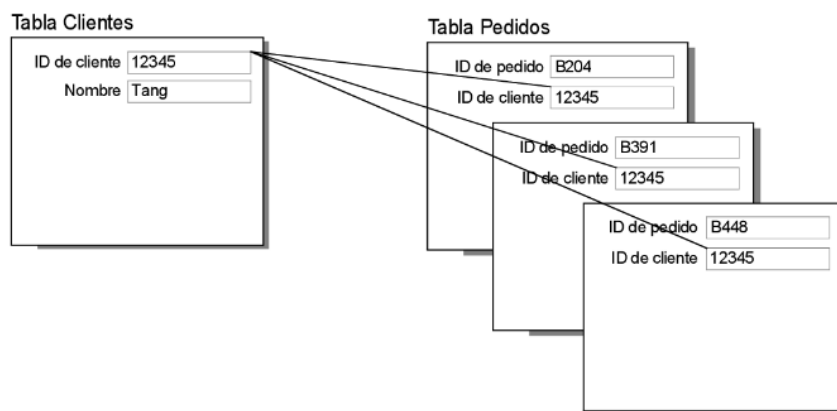
Algunos ejemplos de este tipo de relaciones en bases de datos 1:1 lo encontramos en la asignación de banderas por país, donde cada bandera solo se puede asignar a un único país. Lo mismo ocurre con otros datos como las matrículas de coche (cada coche solo puede tener un único número de matrícula, y cada matrícula solo puede corresponder a un coche en concreto) o el número de serie de un producto (cada producto tiene un número único de serie, y cada número de serie corresponde a un único producto).



Relación uno a varios (1:N)

En este tipo de relación de uno a varios, un registro de una tabla puede enlazarse a varios registros de otra tabla. La primera key está vinculada a varios registros de otra tabla. Esta relación es la más utilizada en las bases de datos relacionales.

Veamos ejemplos de este tipo de relación 1:N:

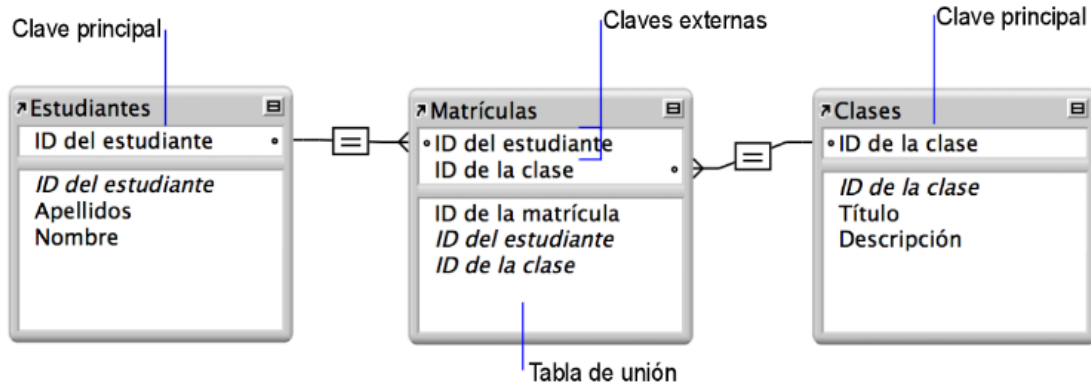


Relaciones varios a varios (N:N)

Cuando varios registros de una tabla pueden relacionarse con varios registros de otra tabla se está ante una relación del tipo varios a varios. Este tipo de relación es la menos habitual en una base de datos relacional.

Un ejemplo de este tipo de tablas N:N lo podemos ver con una tabla de equipo de fútbol y otra de jugadores. Se puede dar el caso de que en un equipo de fútbol juegan varios jugadores y, que además, un jugador haya jugado en varios equipos de fútbol.

Otro ejemplo lo podemos encontrar en una base de datos con una tabla de asignaturas y otra de profesores. Varios profesores pueden dar una misma asignatura, así como un profesor puede impartir varias asignaturas.



Hemos hablado sobre qué es una relación en una base de datos, su utilidad y los tipos que existen. Las relaciones entre entidades o tablas en las bases de datos son esenciales para poder dar lógica y consistencia a los datos, y permiten realizar consultas de datos de forma rápida y precisa.

Las bases de datos relaciones son las más utilizadas al ser más sencillas de construir y gestionar, estableciendo diferentes tipos de relaciones entre los registros de las tablas que las forman.

¿Qué es el lenguaje de consulta estructurada (SQL)?

El SQL es un lenguaje de programación que utilizan casi todas las bases de datos relacionales para consultar, manipular y definir los datos, además de para proporcionar control de acceso. El SQL se desarrolló por primera vez en IBM en la década de 1970 con Oracle como uno de los principales contribuyentes, lo que dio lugar a la implementación del estándar ANSI SQL. El SQL ha propiciado muchas ampliaciones de empresas como IBM, Oracle y Microsoft. Aunque el SQL se sigue utilizando mucho hoy en día, están empezando a aparecer nuevos lenguajes de programación.