



MATERIA:

Introdución a las Bases de Datos

Modulo:

Docente:

Grupo:



Unidad I

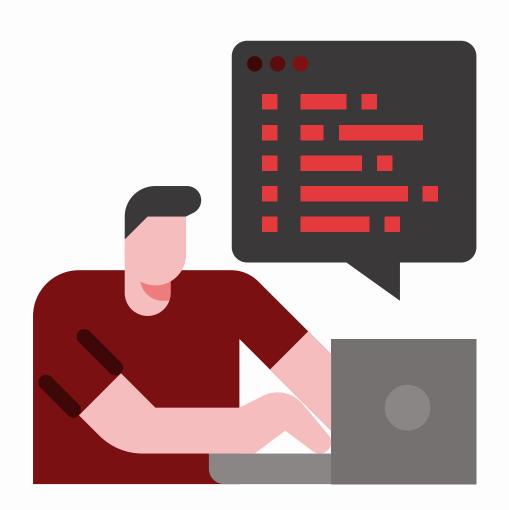
Introducción a la Base de Datos



Introducción a Bases de Datos

- >> Tecnología de los RDBMS
- **DE Evolución de los RDBMS**
- >> Conceptos básicos de las Bases de Datos
- >> Terminología
- **⊳⊳ Modelo de las Bases de Datos**
- **DE Clasificación de las Bases de Datos**





Tecnología de los RDBMS



1.1 Tecnología de los RDBMS

Sistema de gestión de bases de datos

• Un sistema de administración de bases de datos (DBMS) es un software diseñado para almacenar, recuperar y administrar datos en bases de datos. El sistema DBMS más común es el RDBMS. La forma completa de RDBMS es el sistema de gestión de bases de datos relacionales.



1.1 Tecnología de los RDBMS

Necesidad de una base de datos relacional

- Una base de datos relacional está diseñada específicamente para manejar una amplia gama de formatos de datos y procesar consultas de manera eficiente gracias a que se organizan en tablas que se pueden vincular internamente dependiendo de los datos comunes.
- Esto permite que los datos almacenado puedan consultarse y utilizarse de una forma ordenada.



1.1 Tecnología de los RDBMS Como funciona una RDBMS

- Una base de datos relacional almacena datos en forma de varias tablas.
- Una base de datos relacional está diseñada específicamente para manejar una amplia gama de formatos de datos y procesar consultas de manera eficiente gracias a que se organizan en tablas que se pueden vincular internamente dependiendo de los datos comunes.
- Esto permite que los datos almacenado puedan consultarse y utilizarse de una forma ordenada.



1.1 Tecnología de los RDBMS Como funciona un RDBMS

• A continuación, se muestra un ejemplo sencillo de una base de datos. La primera columna contiene el ID de la empresa, que es la clave principal. La clave principal se utiliza para identificar cada registro único en una tabla. Cada columna (o atributo) almacena bits de información y la base de datos asigna un ID de empresa único a cada fila.





1.1 Tecnología de los RDBMS Ventajas de los RDBMS



a) Seguridad de datos mejorada

Las funciones de autorización y control de acceso del software de bases de datos relacionales admiten el cifrado y descifrado avanzados, lo que permite a los administradores de la base de datos gestionar el acceso a los datos almacenados. Esto ofrece importantes beneficios en términos de seguridad.



b) Retener la consistencia de datos

Es más fácil agregar nuevos datos o modificar tablas existentes en un RDBMS mientras se mantiene la coherencia de los datos con el formato existente.



1.1 Tecnología de los RDBMS Ventajas de los RDBMS



c) Mejor flexibilidad y escalabilidad

Un RDBMS ofrece más flexibilidad a la hora de actualizar los datos, ya que las modificaciones solo deben realizarse una vez. Por ejemplo, actualizar los detalles en la tabla principal actualizará automáticamente los archivos relevantes y le evitará la molestia de cambiar varios archivos uno por uno.



d) Facil mantenimiento

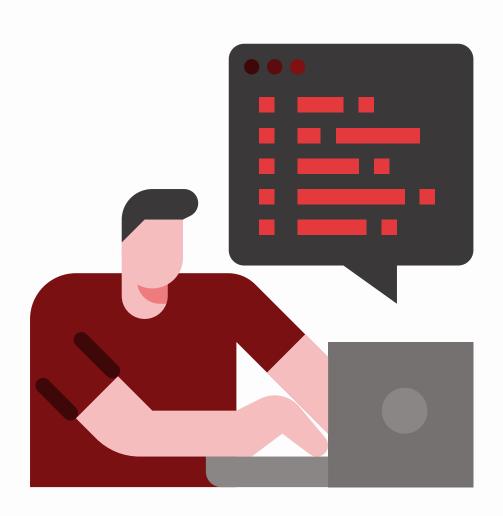
Se consideran de bajo mantenimiento porque los usuarios pueden probar, regular, corregir y realizar copias de seguridad de los datos



e) Riesgo reducido de errores

El software puede verificar fácilmente si hay errores en los datos de diferentes registros.





Evolución de los RDBMS



1.2 Evolución de los RDBMS

Historia de las bases de datos

El término Base de Datos



se usó por primera vez en un simposio celebrado en California, en el año 1963. Su origen se remonta a la antigüedad, cuando ya existían bibliotecas y registros acumulados de hechos y situaciones, de forma escrita y gráfica.

Luego, cuando el volumen de datos e informaciones aumentó, hubo la necesidad de mejorar la capacidad de almacenamiento.

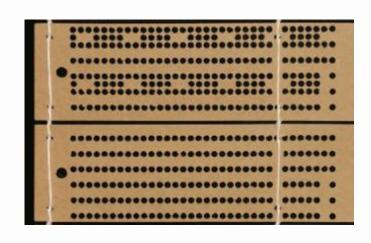
Fueron apareciendo las primeras computadoras, que facilitaron inmensamente el trabajo de guardar gran cantidad de datos, la informática siempre ha estado ligada al concepto de Base de Datos.



1.2 Evolución de los RDBMS

Historia de las bases de datos

Maquina de tarjetas perforadas



En 1884, los censos se realizaban de forma manual, hasta que Herman Hollerith inventó la máquina automática de perforación de tarjetas, que se usó en el censo de los Estados Unidos, mejorando significativamente el proceso de terminación, de siete años a dos años y medio.

En 1950, se cambia a un sistema de lectura secuencial y ordenada. El inglés Oberlin Smith, con este mecanismo, dio inicio a la automatización de la información referente a las nóminas, a través de cintas magnéticas, que a su vez respaldaban dicha información.



1.2 Evolución de los RDBMS Historia de las bases de datos

Evolución de las bases de datos



En 1960, las empresas pudieron adquirir computadoras para facilitar sus gestiones. Las empresas informáticas habían bajado los precios de las mismas, para popularizar el uso de los discos, adelanto muy valioso y útil para esa época, ya que se ubicaba la información de manera directa, sin necesidad de saber la ubicación exacta de los datos. Se convirtió en una opción más rentable para las empresas pequeñas, medianas y grandes.

También se inició la primera generación de bases de datos de red (CODASYL) y las jerárquicas (IMS), que consistían en guardar las estructuras de datos en listas y árboles, además de que permitió crear un estándar en las bases de datos, gracias a los nuevos lenguajes implementados en los sistemas de información.



1.2 Evolución de los RDBMS Historia de las bases de datos

Década de los 70



- Edgar Frank Codd, científico informático inglés, quien definió el modelo relacional. El multimillonario Lawrence "Larry" Ellison, pudo desarrollar el Relational Software System o sistema de datos ORACLE
- Inicialmente no se usó porque tuvo inconvenientes con el rendimiento, no podía competir con las bases de datos jerárquicas y de redes.
- Finalmente IBM desarrolló unas técnicas para construir un sistema de bases de datos relacionales eficientes, las cuales llamó System R; por otro lado Ingres se desarrolló en la UBC en los años de 1974 a 1977



1.2 Evolución de los RDBMS

Historia de las bases de datos

Década de los 80 - Comercialización de sistemas relacionales



- En la década de los años 80', se desarrolló el SQL (Structured Query Language), un lenguaje de consultas que permite consultar, valga la redundancia, con el fin de recuperar información de una base de datos y a su vez, hacer cambios sobre esa misma base, de forma sencilla.
- Permitía analizar gran cantidad de información y especificar varios tipos de operaciones con la misma información
- SQL comenzó a ser el modelo estándar de las industrias, con su base de datos bajo un sistema de tablas (filas y columnas) siendo un éxito comercial.



1.2 Evolución de los RDBMS Historia de las bases de datos

Década de los 90

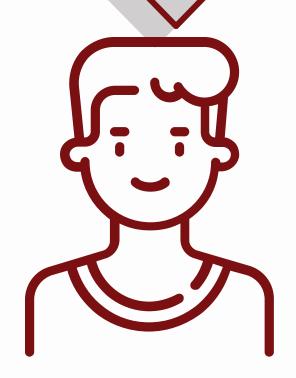
- La investigación en bases de datos se orientó hacia los objetos, con bastante éxito en la gestión de datos complejos.
- Las herramientas para desarrollar las aplicaciones fueron liberadas, incluyendo el desarrollador de Oracle, PowerBuilder, Visual Basic, como también para la productividad personal, las herramientas como ODBC y Excel / Access, también se desarrollaron.
- Con la llegada del Internet la industria de la base de datos creció a un ritmo exponencial. Gran cantidad de usuarios de escritorio comenzaron a usar sistemas cliente-servidor, de bases de datos y acceder a los sistemas informáticos.
- En los 90' también se empezó a modificar la primera publicación de ANSI con respecto al lenguaje SQL, se agregaron nuevas expresiones regulares, triggers, consultas recursivas y características orientadas a objetos.





El principal lenguaje de base de datos y el más utilizado desde que se conoce la programación de gestión, es el Structured Query Language (SQL)

Los gestores de bases de datos, Database Management System o DBMS (SGBD) son programas que permiten almacenar y luego acceder a los datos de forma estructurada y rápida. Las aplicaciones más usadas son para gestiones de empresas e instituciones públicas, así como en entornos científicos, para almacenar la información experimental.





1.2 Evolución de los RDBMS Historia de las bases de datos













1.2 Evolución de los RDBMS

Historia de las bases de datos

Actualidad de las bases de datos

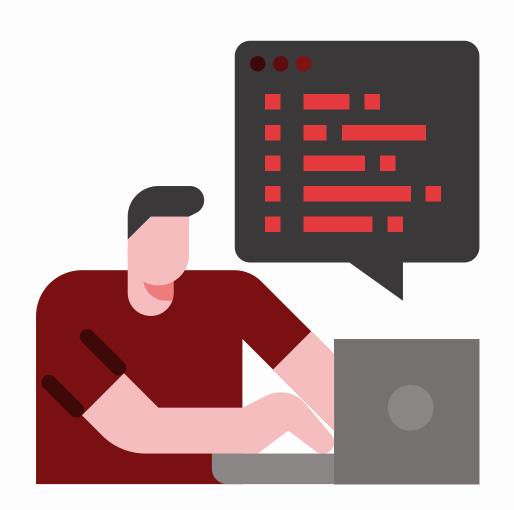


Hoy día, las compañías que dominan el mercado de las bases de datos son IBM, Microsoft y Oracle. Mientras que en lo que respecta a internet, el dominante es Google, generando gran cantidad de información.

Aun cuando también existe gran cantidad de softwares que permiten crear y manipular bases de datos con gran facilidad, como LINQ de Microsoft, que se asemeja un poco a SQL en cuanto a las consultas nativas en los lenguajes de la plataforma .NET.

Cada dia las bases de datos toman tamaños cada vez mas grandes y de real importancia en aplicaciones de BigData, Ciencia de Datos, Machine Learning, IA, etc.



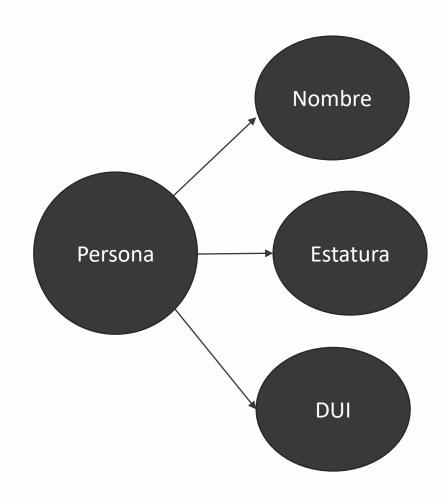


Conceptos Basicos



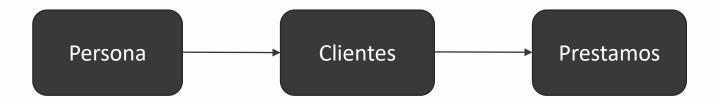
Entidad

Una entidad es una «cosa» u «objeto» en el mundo real que es distinguible de todos los demás objetos. Por ejemplo, cada persona en un desarrollo es una entidad. Una entidad tiene un conjunto de propiedades, y los valores para algún conjunto de propiedades pueden identificar una entidad de forma unívoca.



Conjunto de entidades

Es un conjunto de entidades del mismo tipo que comparten las mismas propiedades, o atributos. El conjunto de todas las personas que son clientes en un banco dado, por ejemplo, se pueden definir como el conjunto de entidades cliente. Análogamente, el conjunto de entidades préstamo podría representar el conjunto de todos los préstamos concedidos por un banco particular





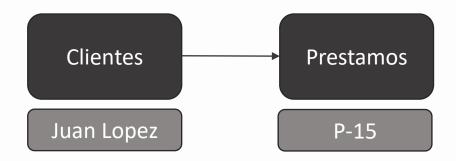
Atributos

Los atributos describen propiedades que posee cada miembro de un conjunto de entidades. La designación de un atributo para un conjunto de entidades expresa que la base de datos almacena información similar concerniente a cada entidad del conjunto de entidades; sin embargo, cada entidad puede tener su propio valor para cada atributo.

Cada entidad tiene un valor para cada uno de sus atributos. Por ejemplo, una entidad cliente en concreto puede tener el valor 34546 para id-cliente, el valor "José" para nombre-cliente, el valor "Independencia" para callecliente y el valor "San Salvador" para ciudad-cliente.

Relacion

Una relación es una asociación entre diferentes entidades. Por ejemplo, se puede definir una relación que asocie al cliente Juan López con el préstamo P-15. Esta relación especifica que Juan López es un cliente con el préstamo número P-15.



Tabla

Las tablas son objetos de base de datos que contienen todos sus datos. En las tablas, los datos se organizan con arreglo a un formato de filas y columnas, similar al de una hoja de cálculo. Cada fila representa un registro único y cada columna un campo dentro del registro.

Por ejemplo, en una tabla que contiene los datos de los empleados de una compañía puede haber una fila para cada empleado y distintas columnas en las que figuren detalles de los mismos, como el número de empleado, el nombre, la dirección, el puesto que ocupa y su número de teléfono particular.

Id_empleado	Nombre	Apellido	Cargo
001	Juan	Lopez	Director
002	Pedro	Perez	SubDirector

Registros

Son la unidad básica que describe a un objeto o una transacción. Las tablas están formadas por un conjunto de ellos.

Campos o Columnas

El registro está formado por uno o más campos, que son la unidad básica de información, y son una propiedad específica de un objeto o transacción.

Id_transaccion	Servicio	Hora
12895	Α	06:00
33254	В	09:00

Dato

Es la intersección de un campo y un registro. Es toda aquella información que puede almacenarse en una base de datos.

Tipo de dato

Dependiendo de la columna, cada una de ellas tendrá definido un tipo de dato específico que indicará la forma de almacenar, las características y los límites.

Así, hay distintos tipos: los hay numérico, alfanumérico o de tipo cadena, de tipo fecha, etc.

Cada motor de base de datos define distintos tipos de datos, y algunos son específicos de cada una de ellas.

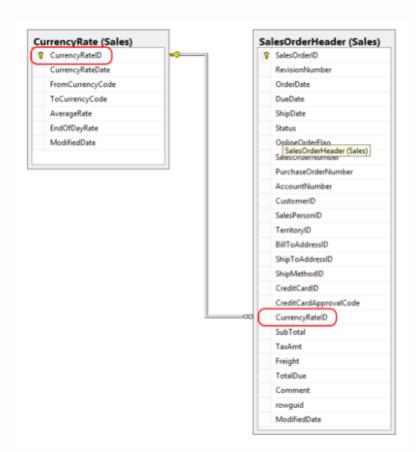


Clave Primaria

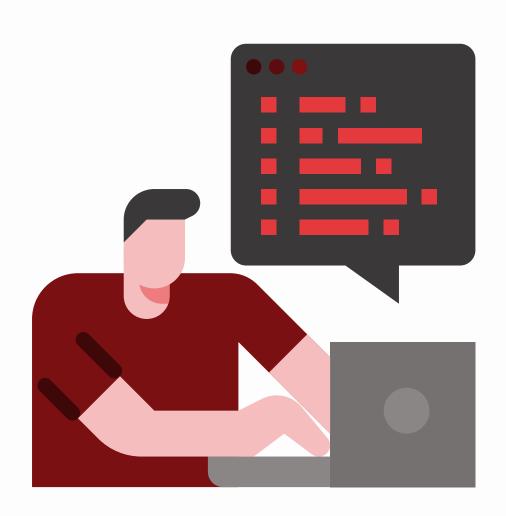
Es una columna o un conjunto de columnas en una tabla cuyos valores identifican de forma exclusiva una fila de la tabla. Una base de datos relacional está diseñada para imponer la exclusividad de las claves primarias permitiendo que haya sólo una fila con un valor de clave primaria específico en una tabla.

Clave foránea

Es una columna o un conjunto de columnas en una tabla cuyos valores corresponden a los valores de la clave primaria de otra tabla. Para poder añadir una fila con un valor de clave foránea específico, debe existir una fila en la tabla relacionada con el mismo valor de clave primaria.







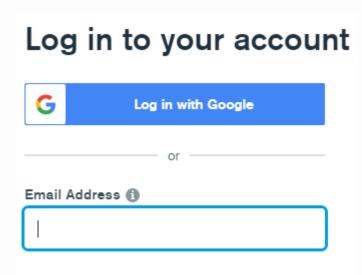
Terminología



Formularios

Un formulario es un objeto de base de datos que se puede usar para escribir, modificar o mostrar los datos de una tabla o consulta. Los formularios se pueden usar para controlar el acceso a los datos, como qué campos o filas de datos se van a mostrar. Por ejemplo, puede que algunos usuarios necesiten ver sólo algunos de los campos de una tabla que contiene numerosos campos.

Permite definir cómo se agrupan y se ordenan los datos, y usar campos de más de una tabla o consulta siempre y cuando defina con antelación las relaciones entre las tablas y consultas.



Consulta(Query)

En bases de datos, una consulta es el método para acceder a los datos en las bases de datos. Con las consultas se puede modificar, borrar, mostrar y agregar datos en una base de datos. Para esto se utiliza un lenguaje de consultas. El lenguaje de consultas a base de datos más utilizado es el SQL.

SELECT * FROM table_name;



Informe

Los informes son resúmenes de la información de la base de datos que puedes consultar en pantalla o mediante el uso de una impresora. Los informes pueden ser una simple lista de los registros de una tabla, o los más complejos que incluyen varias tablas, resumen de totales y tablas o gráficos. Determina qué campos y tablas contienen la información que necesitas, y utilizar el asistente de informes o la ventana de diseño de informes para crear para crearlos.





Macro

Una macro es un conjunto de comandos que se almacena en un lugar especial de Excel de manera que están siempre disponibles cuando los necesites ejecutar.

Por ejemplo, si todas las mañanas creas un reporte de ventas y en ese reporte siempre das el mismo formato a los textos, se podría crear una macro para que lo haga automáticamente por ti. Las macros se utilizan principalmente para eliminar la necesidad de repetir los pasos de aquellas tareas que realizas una y otra vez.





Procedimiento Almacenado

Un procedimiento almacenado (stored procedure en inglés) es un programa (o procedimiento) almacenado físicamente en una base de datos. Su implementación varía de un gestor de bases de datos a otro. La ventaja de un procedimiento almacenado es que al ser ejecutado, en respuesta a una petición de usuario, es ejecutado directamente en el motor de bases de datos, el cual usualmente corre en un servidor separado. Como tal, posee acceso directo a los datos que necesita manipular y sólo necesita enviar sus resultados de regreso al usuario, deshaciéndose de la sobrecarga resultante de comunicar grandes cantidades de datos salientes y entrantes.

```
☐ CREATE PROCEDURE #spu_ObtenerTablasColumnas

AS

☐ SELECT t.TABLE_SCHEMA, t.TABLE_NAME, COUNT(1) AS NUMERO_COLUMNAS

FROM INFORMATION_SCHEMA.TABLES t

INNER JOIN INFORMATION_SCHEMA.COLUMNS c

on t.TABLE_SCHEMA = c.TABLE_SCHEMA

and t.TABLE_NAME = c.TABLE_NAME

GROUP BY t.TABLE_SCHEMA, t.TABLE_NAME

ORDER BY 1,2

GO

EXEC #spu_ObtenerTablasColumnas
```



Trigger

Un trigger o disparador es un script que se usa en lenguaje de programación SQL, en especial en bases de datos como MySQL o PostgreSQL.

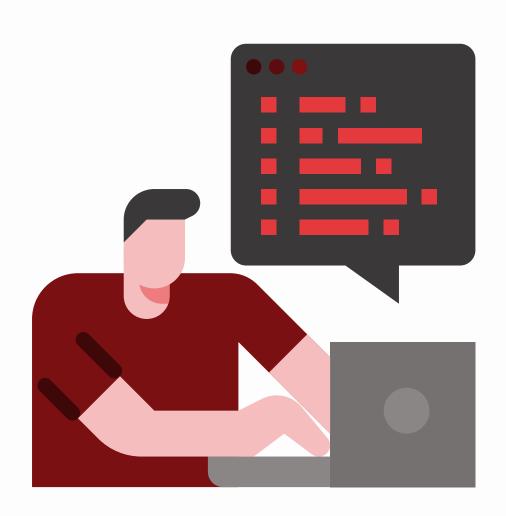
Consiste en una serie de reglas predefinidas que se asocian a una tabla. Estas reglas se aplican a la base de datos cuando se realizan determinadas operaciones en la tabla, por ejemplo, al añadir, actualizar o eliminar registros.

Dicho de otra manera, el trigger desencadena determinadas acciones de forma automática en las tablas de la base de datos cuando se insertan, modifican y se añaden nuevos datos.

Índice

Un índice es una estructura diferente dentro de la base de datos; creado con el comando create index. Requiere su propio espacio en disco y contiene una copia de los datos de la tabla. Eso significa que un índice es una redundancia. Crear un índice no cambia los datos de la tabla; solamente establece una nueva estructura de datos que hace referencia a la tabla. De hecho, un índice de base de datos se parece mucho a un índice de un libro: ocupa su propio espacio, es redundante y hace referencia a la información actual almacenada en otro lugar.



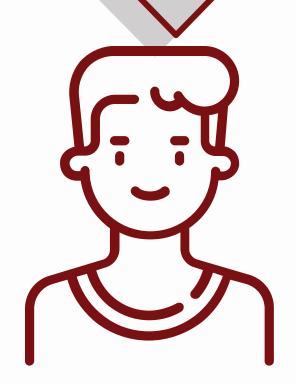


Modelo de las bases de datos



Un modelo de base de datos es la estructura lógica que adopta la base de base datos, incluyendo las relaciones y limitaciones que determinan cómo se almacenan y organizan y cómo se accede a los datos. Así mismo, un modelo de base de datos también define qué tipo de operaciones se pueden realizar con los datos, es decir, que también determina cómo se manipulan los mismos, proporcionando también la base sobre la que se diseña el lenguaje de consultas.

Que son los modelos de bases de datos y que tipos existen.





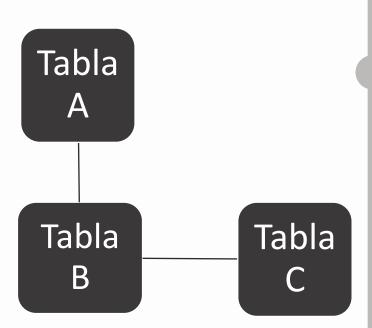
Modelo de base de datos relacional

Es uno de los más comunes y emplean las bases de datos relacionales y ordena los datos en tablas (relaciones) compuestas por columnas y filas.

Cada columna alberga un atributo de la entidad (nombre, dirección, fecha de nacimiento...); a los atributos de una relación se los llama dominio. Escogiendo un atributo en concreto o una combinación de varios tenemos una clave primaria, a la que se puede hacer referencia en otras tablas, en las que será un clave externa.

En cada fila (tupla) se incluyen datos sobre una instancia específica de la entidad (por ejemplo, un cliente específico).

Además, el modelo también representa el tipo de relaciones entre las tablas, que pueden ser uno a uno, uno a muchos o muchos a muchos.

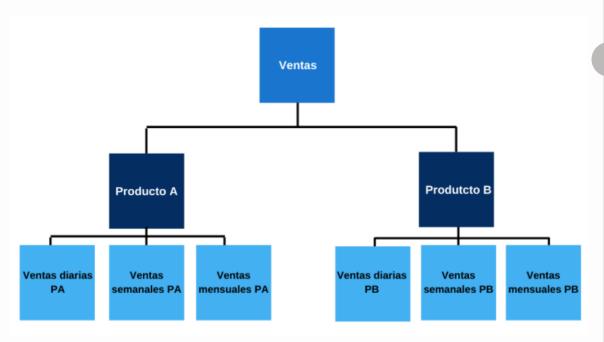




Modelo jerárquico

Si vamos a emplear una base datos jerárquica, el modelo de datos que emplearemos será el jerárquico, que se caracteriza por presentar los datos en una estructura de árbol invertido, donde cada registro tiene un único nodo raíz, del que surgen otros nodos (registros); los nodos en un mismo nivel son nodos padre, cada nodo padre tiene el mismo nodo raíz, y puede tener nodos hijos, pero los nodos hijos solo pueden tener un nodo padre. Este modelo se emplea poco actualmente.

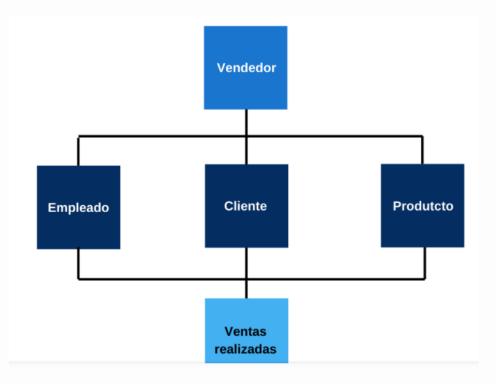
En este modelo, los registros de un mismo nivel se clasifican en un orden específico.





Modelo de red

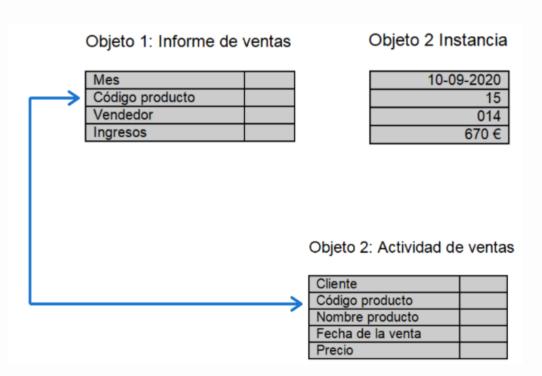
El modelo en red de base de datos parte del modelo jerárquico, pero aquí se permiten las relaciones de uno a muchos o de muchos a muchos entre registros vinculados, teniendo registros principales múltiples. El modelo se crea a través de conjuntos de registros relacionados; cada uno de estos conjuntos consiste en un registro propietario o principal y uno o más registros miembros o secundarios. Además, un registro puede ser miembro o secundario en diferentes conjuntos. Es decir, que en este modelo se permite que los nodos hijos tengan más de uno nodo padre, de manera que se pueden representar relaciones más complejas.





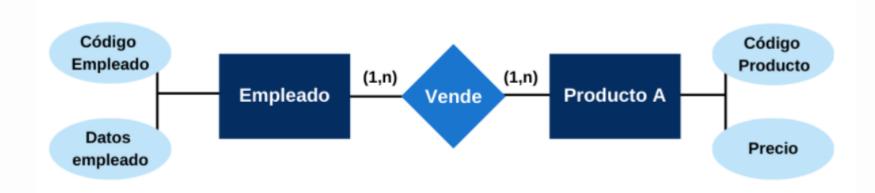
Modelo Orientado a Objetos

El modelo de la base de datos orientada a objetos define la base de datos como una colección de objetos utilizados en la programación orientada a objetos (es decir, que emplear lenguajes como C++ o Java, por ejemplo). Este modelo de base de datos utiliza tablas también, pero no solo se limita a ellas y permite almacenar información muy detallada sobre cada objeto. Los objetos se dotan de un conjunto de características propias, que a su vez les diferencian de objetos similares. Los objetos similares pueden agruparse en una clase y cada objeto de esta es una instancia. Las clases intercambian datos entre sí a través métodos (mensajes).



Modelo entidad-relación

El modelo entidad-relación es básicamente el paso previo a uno modelo de bases datos relacional, puesto que se trata de un diagrama elaborado a través de unos elementos básicos y su relación entre ellos.





Modelo plano

El modelo de bases de datos plano, los datos se estructura en dos dimensiones (de hay lo de estructura plana), en la que todos los objetos en una columna concreta tienen valores del mismo tipo y todos los objetos de la misma fila están relacionados entre ellos.

Por ejemplo, en una base de datos que recoja solo el nombre de usuario y la contraseña, cada fila recogerá el nombre y la contraseña correspondiente para cada usuario.

Este modelo tabular de base de datos es el precursor del modelo relacional.

Usuario	Contraseña
0001	XXXX
0002	уууу
0003	ZZZZ
0004	VVVV



Modelo multidimensional

El modelo de base de datos multidimensional está pensado para la creación de aplicaciones específicas OLAP (procesamiento analítico en línea). Este modelo de base de datos se puede visualizar como un cubo de datos en el que se representan diferentes dimensiones de los datos disponibles; las dimensiones de los cubos se corresponden con la de la tabla y el valor almacenado en cada celda equivale al valor de la métrica.

Modelo semiestructurado

Cuando los datos no encajan en el formato de tablas, filas y columnas, sino que se organizan mediante etiquetas con las que se los puede agrupar y crear jerarquías (como ocurre con los datos de una página web), estamos hablando de datos semiestructurados.

Para ordenador estos datos en una base de datos, podemos recurrir al modelo semiestructurado. En este modelo no hay separación entre los datos y el esquema y resulta útil para describir sistemas o describir interacciones entre bases de datos no que se acogen al mismo esquema.



Modelo de contexto

El modelo de contexto se puede usar cuando se necesita incorporar elementos de otros modelos de bases de datos. Puede adoptar elementos de modelos orientados a objetos, semiestructurados y en red.

Modelo asociativo

En modelo asociativo los datos se dividen en entidad y asociación, de manera que una entidad es todo lo que existe de manera independiente y una asociación es algo que solo existen en relación a algo más. Así, los datos se estructuran en dos grupos:

- Elementos, donde cada uno tiene un identificador único, un nombre y un tipo.
- Enlaces, donde cada tiene un identificador único e identificadores únicos de una fuente, verbo u objeto. La información almacenada está relacionada con la fuente y cada uno de los tres identificadores pueden hacer referencia a un enlace o a un elemento.



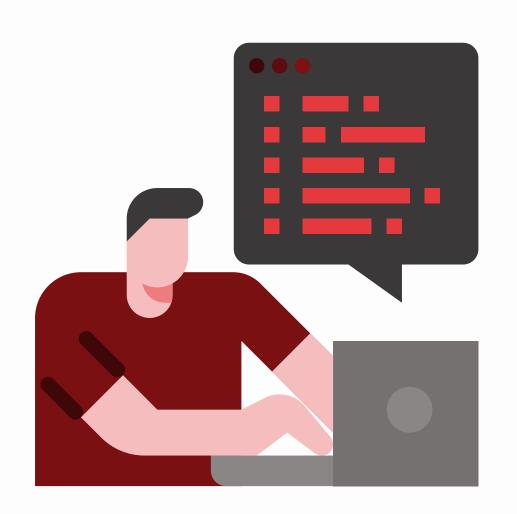
Modelos de bases de datos NoSQL

también llamadas bases de datos no relacionales, el modelo de base de datos orientado a objetos, pero existen otros:

- Modelo de base de datos gráfico, similar al de red, pero más flexible, puesto que permite que cualquier nodo se pueda conectar a cualquier otro.
- Modelo multivalor, que nace del modelo relacional, pero en el que los atributos pueden contener una lista de datos en vez de un solo punto de datos.
- Modelo de documentos, empleado para almacenar y administrar documentos o datos semiestructurados, en vez de datos atómicos (como hacen las bases relacionales).







Clasificación de las bases de datos



1.6 Clasificación de las bases de datos

Según la flexibilidad de modificación

La primera clasificación de bases de datos depende de la manera en la que se ordenan los datos.

Bases de datos estáticas:

las BBDD estáticas están diseñadas para la lectura de datos. En otras palabras, sólo almacenan y registran los datos. Luego se pueden analizar para comprender su comportamiento a lo largo del tiempo. Son especialmente utilizadas para implementar proyecciones estadísticas y orientar procesos de tomas de decisiones en el ámbito empresarial.

Bases de datos dinámicas:

las BBDD dinámicas son, por el contrario, modificables con el paso del tiempo. Así, los datos pueden actualizarse, editarse y eliminarse. Por ejemplo, muchas tiendas varían sus inventarios y los precios de sus productos según la temporada, por lo que una BBDD dinámica es ideal.



1.6 Clasificación de las bases de datos

Según el contenido

La segunda clasificación de bases de datos depende de la prioridad del contenido a analizar.

Bases de datos bibliográficas

Las BBDD bibliográficas son registros que ayudan a clasificar varios campos de datos. Generalmente, dichos campos se pueden consultar de modo separado o conjunto. Un claro ejemplo puede ser la información sobre un libro: autor, año de publicación, editorial, etc.

Bases de datos de texto completo

Entre los tipos de bases de datos, las de texto completo son especialmente útiles, puesto que permiten buscar términos específicos, palabras claves y las diversas opciones de una base de datos bibliográfica, además de consultar el texto íntegro almacenado. Son idóneas para trabajos académicos y de investigación.



1.6 Clasificación de las bases de datos

Directorios:

Los directorios son BBDD usadas por la mayoría de la gente casi a diario sin darse cuenta. Un ejemplo claro puede ser la agenda de contactos de nuestros teléfonos móviles, donde se almacena multitud de información como:

- Nombres y direcciones
- Número telefónico y direcciones de email
- Datos de facturación, códigos postales