**Clase 01 - Introducción a Bases de datos**

**Concepto Base de Datos(DB Data base)**

Conjunto de información, usualmente perteneciente a un mismo contexto, que es almacenada sistemáticamente para su posterior uso.

**Base de Datos Analógicas Vs. Digitales**

En el **mundo analógico**, una biblioteca es la representación más fehaciente de una Base de datos. Un placard también puede ser considerado como una base o almacén de datos; dado que contiene compartimentos y estos, a su vez, contienen prendas de vestir, calzado, ropa interior, prendas de abrigo, accesorios, etc; todo organizado de manera sistemática.

Un mueble, ya sea de cocina y de living comedor, también puede ser considerado como un almacén de datos u objetos, porque todo lo que guarda o contiene en sus estantes o compartimentos, también está organizado de una manera sistemática.

Un archivero u organizador de carpetas, usualmente encontrados en los sectores administrativos de una empresa, también se consideran como base de datos que almacenan información importante de clientes, proveedores, facturas de compra, facturas de ventas, registros de empleados, libros contables. Todo esto, siempre organizado, de forma sistemática para que acceder a toda esa información, sea de la forma más fácil e intuitiva posible.

Mientras que en el **mundo del software**, tenemos al menos dos definiciones para esto; los discos rígidos o de estado sólido son considerados también como un sistema de almacenamiento del tipo DB por cómo estructuran la información.Y por supuesto, las aplicaciones de software de bases de datos construidas para tal fin.

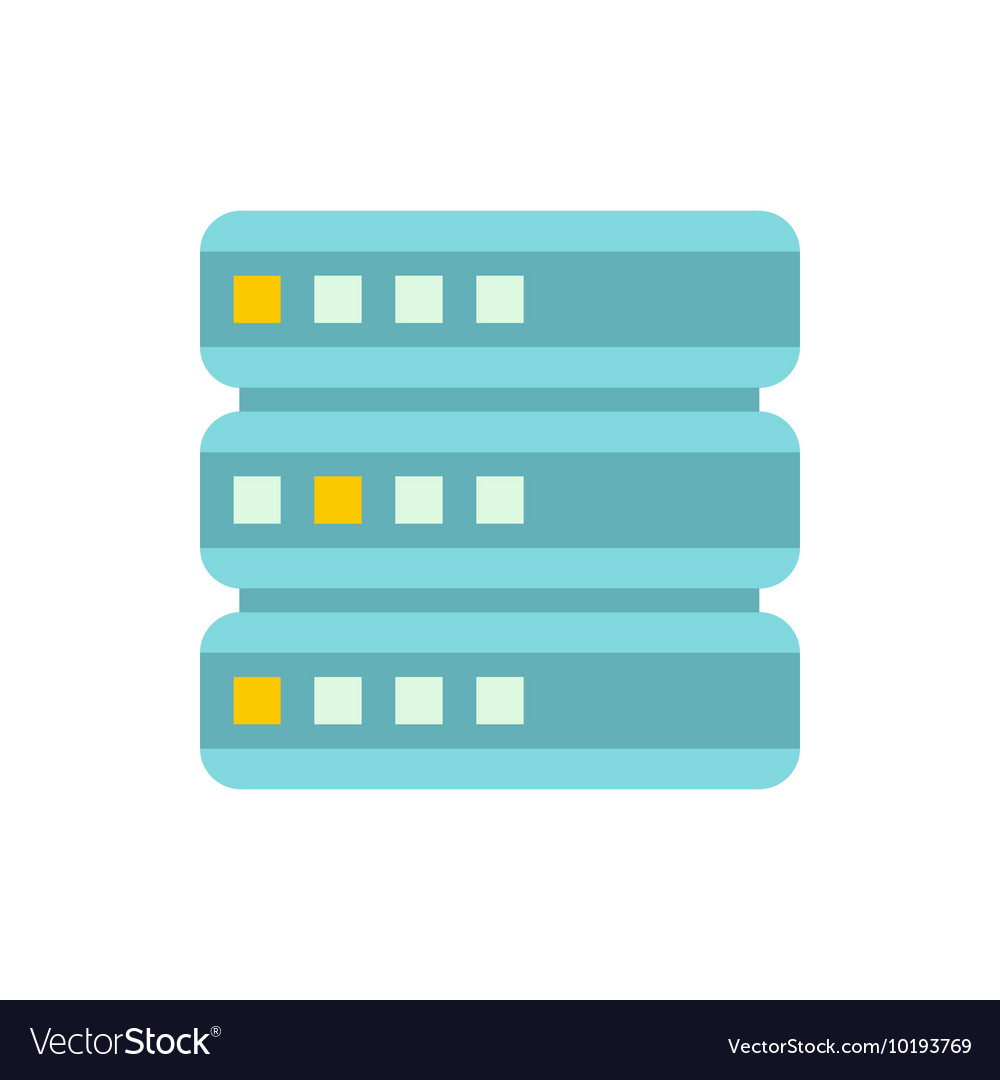
Podemos mencionar al disco rígido, disco de estado sólido, o cualquier otro medio de almacenamiento de información digital, como una base o almacén de datos.

Pensando que el disco o soporte físico representa a la base de datos en sí. En su interior, encontramos las carpetas y archivos, quienes mantienen una estructura lógica y ordenada, sistematizando la información que guardan (gráfica y/o textual).

Es la misma filosofía que se heredó del mundo físico, para almacenar cualquier tipo de elemento perteneciente a un mismo contexto.

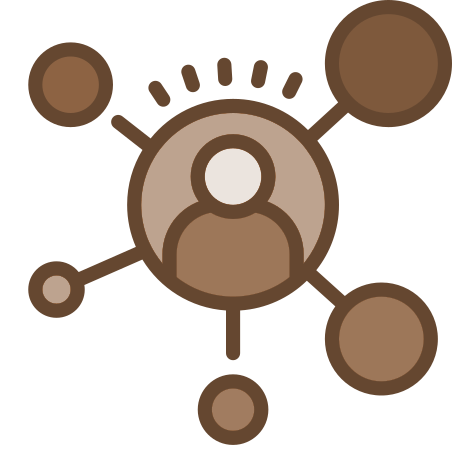
Las aplicaciones de software de bases de datos, almacenan la información bajo cierta lógica, casi siempre estructurada, para acceder a los datos de forma homogénea y que los mismos sean visualizados de la manera más simple y clara posible.

**Usos Más Frecuentes**

**Almacenamiento de datos corporativos:** fue el primer nicho de importancia donde las DB, cumplieron un importantísimo papel. Las bases de datos asociadas a sistemas de administración del tipo contable; como ser SAP, Tango Software, JD Edwards, People Soft, entre otros tantos; son las primeras aplicaciones que más aprovecharon a las DB para no solo almacenar la información que estas plataformas generan en cada corporación, sino también para operar con cada dato generado, recuperar y visualizar a la misma de una manera óptima, gracias al lenguaje SQL que potencia cada una de estas aplicaciones de software.

**La web:** es otro segmento donde las bases de datos cumplen un rol destacable para proveernos en tiempo real la información que buscamos.

Cada buscador integrado a Internet, tiene una DB de trasfondo donde se aloja la información indexada de cada sitio web. Estas DB son de las denominadas “de misión crítica”, porque deben responder no solo con precisión en lo que el usuario solicita, sino que también atienden la demanda de millones de personas concurrentes durante las 24 horas del día.

**Redes Sociales:** El auge de las redes sociales fue posible gracias a una o varias bases de datos potentes que puedan responder de igual forma a una cantidad de accesos concurrentes, de manera eficaz. Con el tiempo, las empresas de redes sociales comenzaron a aprovechar todos esos datos para poder vender publicidad “netamente enfocada en cada nicho de usuario”.

**Análisis de datos:** Del punto anterior nació el análisis de datos basado en el comportamiento del usuario (consumidor). Hoy este análisis no solo enfoca la publicidad en las personas correctas (potencial cliente), sino que también se ocupan de predecir.

La inteligencia artificial, machine learning y IoT, entre otros segmentos asociados, permite entrenar a las computadoras mediante la ingesta de datos constante, y así poder resolver dilemas, automatizar situaciones de la vida, o mejorar procesos de todo tipo de segmento, gracias a la generación constante de información atomizada.

**Historia**

LA DÉCADA DEL 60  
Antes del invento del modelo de base de datos relacional, más precisamente a inicios de la década del 60’, el término base de datos comienza a aparecer en libros de computación.

En esa época, los Mainframe ganaban protagonismo y comenzaban a ser utilizados en universidades, centros de salud, centros gubernamentales, y centros militares. Y toda la información que generaban y procesaban, se almacenaba en tarjetas perforadas y/o en cintas magnéticas similares a la de casetes, pero bastante más grandes.

UN VIAJE EN EL TIEMPO  
Nueva York, 1884: Este podría ser el año y el lugar exacto donde podemos encontrar el origen de las actuales bases de datos.

Desde entonces y hasta ahora las sociedades han tenido la necesidad de conservar datos e información. Y aquí [Herman Hollerith](https://www.biografiasyvidas.com/biografia/h/hollerith.htm) tenía mucho que decir.

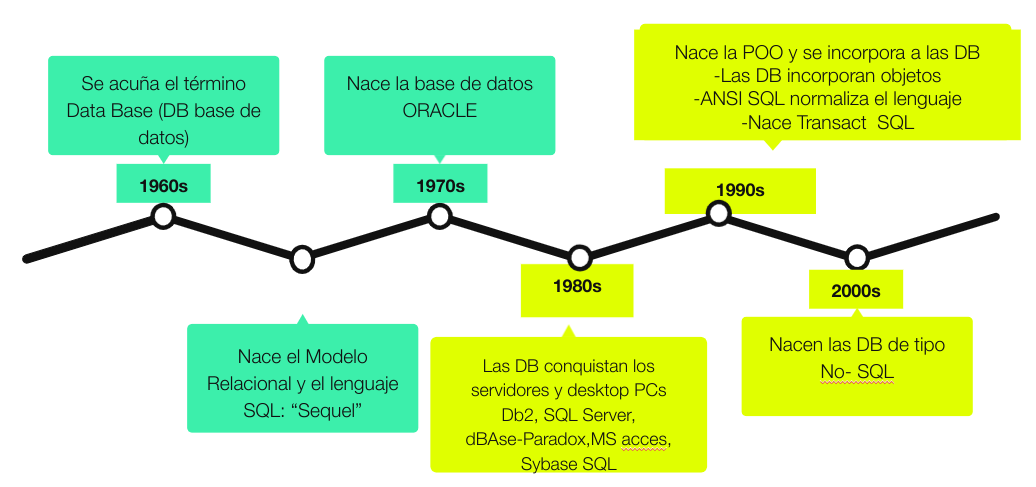
Hollerith fue el creador de la máquina automática de tarjetas perforadas, considerada como el origen de las bases de datos actuales. Sin saberlo Hollerith se convertiría en el primer ingeniero estadístico de la historia. Y es que durante esta época aún tanto en la vieja Europa como en los incipientes Estados Unidos, los censos se realizaban de forma manual. Hollerith comenzó a trabajar en el diseño de una máquina tabuladora o censadora, basada en tarjetas perforadas.

Lo curioso de todo esto es que nadie más avanzó en ninguna dirección en décadas posteriores pese a los avances técnicos que ofrecía el lógico progreso de la sociedad. Tuvieron que pasar más de sesenta años hasta que en 1950 se diera el siguiente paso con las cintas magnéticas; un tipo de medio o soporte de almacenamiento de información que grababa en pistas sobre una banda plástica con un material magnetizado que generalmente era de óxido de hierro o algún cromato. Este invento sirvió para suplir las necesidades de información de la industria. Con ellas se empezó a automatizar, por fin, la información. Las bases de datos actuales estaban cada vez más cerca.

En la década de los 70’, [Edgar Frank Codd](https://www.ecured.cu/Edgar_Frank_Codd) habla sobre base de datos relacionales y se desarrolla la base de datos Oracle.

En los 80, estas bases de datos relacionales logran posicionarse en el mercado de base de datos con sus funcionales sistemas de tablas, filas, columnas; que derivarían en los 90’ con el desarrollo de herramientas como Excel o Access de Microsoft. Una nueva generación de bases de datos que han cambiado la manera de trabajar de prácticamente todas las empresas y que hoy en día siguen evolucionando en un campo tan inabarcable como sorprendente.

Extraído de: <https://www.camerdata.es/blog/historia-bases-datos/>



PARA INVESTIGAR LA HISTORIA LLEVADA AL CINE  
Si quieren ver referencias a los mainframes de época mencionados, la película Talentos Ocultos muestra la importancia de los mainframes, el nacimiento del lenguaje de programación COBOL quien almacenaba o imprimía la información en tarjetas perforadas. Otra serie de referencia a mainframes y la importancia del almacenamiento de datos es HALT AND CATCH FIRE, al menos en sus primeras tres temporadas.

Pueden buscar fragmentos de esto en Youtube.

**Tipos de DB**

**DB Distribuidas:** En comparación con la idea de la base de datos centralizada; existen entradas de la base de datos general y la información recopilada de las computadoras locales. Los datos no son accesibles en una sola ubicación y se distribuyen a varios sitios de la empresa. Estos sitios están conectados entre sí a través de enlaces de comunicación que permiten el acceso a los datos distribuidos.

**DB Relacionales:** Dichas bases de datos se clasifican mediante un conjunto de tablas, en las que los datos entran en una clasificación predefinida. La tabla está formada por filas y columnas con entrada de datos para una determinada categoría y filas; con el ejemplo de los datos identificados por la categoría.

El [lenguaje de consulta estructurado](https://www.educba.com/what-is-sql/) es la interfaz estándar de un programa de aplicación y usuario de base de datos de relaciones.

**DB Orientadas a Objetos:** es una colección de bases de datos relacionales basada en objetos. Hay diferentes elementos, como java, C ++, que se pueden guardar en una base de datos relacional utilizando lenguajes de programación orientados a objetos, pero las bases de datos orientadas a objetos son adecuadas para estos componentes.

Se organizará una base de datos orientada a objetos en lugar de acciones en torno a objetos; y datos en lugar de lógica. A diferencia de un valor alfanumérico, por ejemplo, un registro multimedia en una base de datos relacional puede ser un objeto de datos definible.

**DB en la nube:** Hoy en día, los datos se almacenan en una nube pública, una nube híbrida o una nube privada, también conocida como entorno virtual. Una base de datos en la nube es una base de datos automatizada o integrada para dicho entorno virtualizado.

Un servicio en la nube ofrece varias ventajas, incluida la capacidad de pagar por la capacidad de almacenamiento y el ancho de banda del usuario y proporciona escalabilidad a pedido, así como alta disponibilidad.

**DB Centralizada:** Los datos se almacenan de forma centralizada y los usuarios de varias ubicaciones pueden acceder a estos datos. Esta base de datos incluye procesos de contratación que ayudan a los usuarios incluso desde una ubicación remota a acceder a los datos. Para la verificación y validación de los usuarios finales, se aplican varios tipos de procedimientos de autenticación, y los procesos de aplicación que mantienen un seguimiento y registro de la utilización de datos también proporcionan números de registro.

**DB NoSQL:** Se utilizan para grandes conjuntos de datos. Existen ciertos problemas de rendimiento de big data que son manejados de manera efectiva por bases de datos relacionales, y las [bases de datos NoSQL](https://www.educba.com/what-is-nosql-database/) pueden abordar fácilmente dichos problemas. El análisis de información no estructurada de gran tamaño se puede realizar de manera muy eficiente en varios servidores virtuales en la nube.

**Servidor de DB LOCAL vs. REMOTO**

Al Sistema de Gestión de DB (**SGBD**) (o Database Management System **DBMS** ) nos podemos conectar de manera local, es decir dentro del equipo donde reside o de manera remota, en otro u otros equipos.

Cuando se utiliza conexión remota se dice que se tiene un servidor de DB, se utiliza internet y un puerto habilitado para establecer la conexión.

**SISTEMA DE GESTIÓN DE DB**

SGBD o DBMS se le denomina al conjunto de componentes de software que permiten realizar todo tipo de gestión y operaciones sobre la información que la base de datos almacena, como también sobre la base de datos en sí.

**SISTEMA DE GESTIÓN DE BASE DE DATOS**

**Interfaz** **gráfica** que nos permite acceder visualmente a todos los objetos de una base de datos, y operar sobre ellos.

Existen herramientas que nos permiten hacer lo mismo que un **SGBD**, pero directamente desde la línea de comandos, o ventana Terminal.

***OPERACIONES SOBRE UNA DB***

Las operaciones más frecuentes sobre una base de datos, se denominan bajo el término **C.R.U.D**.  
(**Crear**, **Leer**, **Actualizar** y **Eliminar**)

Estas operaciones se pueden realizar sobre los datos almacenados y también sobre cualquier objeto que compone la base de datos

***BASES DE DATOS SQL***

Son DB de tipo Relacionales y estructurales. Su nombre SQL proviene del término (**Structured Query Language**), en español: **Lenguaje de Consulta Estructurado.**

Se describe como un **lenguaje declarativo.**

Pasó a ser el estándar del Instituto Nacional Estadounidense de Estándares (**ANSI) en 1986** y de la Organización Internacional de Normalización **(ISO) en 1987.**

**LENGUAJES Y SUBLENGUAJES**

**DML**:Utilizando instrucciones de SQL, **permite a los usuarios introducir datos para posteriormente realizar tareas de consultas o modificación** de los datos que contienen las Bases de Datos.

**DDL:** Este lenguaje **permite a los programadores** de un sistema gestor de base de datos, **definir las estructuras que almacenarán los datos** así como los procedimientos o funciones que permitan consultarlos.

**DCL:** Estos comandos **permiten al Administrador del sistema gestor de base de datos, controlar el acceso a los objetos**, es decir, podemos otorgar o denegar permisos a uno o más roles para realizar determinadas tareas.

**TCL:** son sentencias SQL que permiten manejar transacciones en una DB

**DB NoSQL**

**Versión inicial del término**: DB no relacional. No existía posibilidad de usar SQL como lenguaje de consulta.

**Versión actual del término**: No **solo** SQL. Existen implementaciones que permiten **versiones restringidas** de SQL en la expresión de consultas.

Este tipo de bases de datos fueron popularizadas por Google, Facebook, Amazon y Twitter, entre otros gigantes de la industria. Actualmente, los sistemas basados en nube proveen como alternativa a las comerciales SQL, sus DB NoSQL, las cuales son más fáciles de manipular, sobre todo en el momento en que entran en acción las [APIs](https://ladatacuenta.com/2018/03/26/apis-y-json-que-son-y-para-que-sirven-cuando-trabajas-con-datos/) o [Servicios Web](https://es.wikipedia.org/wiki/Servicio_web), que deben proveer de datos a aplicaciones propias o de terceros.

Algunas DB NoSQl tienen una estructura de clave-valor y entregan la información mediante API en el formato de transporte correcto. En cambio, las bases de datos SQL, necesitan de un conector que se ocupe de transformar el conjunto de registros a proveer el [formato de transporte JSON](https://es.wikipedia.org/wiki/JSON), o [formato XML](https://es.wikipedia.org/wiki/Extensible_Markup_Language), o cualquier otro que se utilice, para recién entregar esta información a la aplicación que la solicita.

TIPOS DE DB NoSQL

**Key-Value:** Bases de datos clave-valor, son las bases de datos NoSQL más simples. Cada elemento de la base de datos se almacena como un nombre de atributo (o «clave»), junto con su valor. Ejemplos de almacenes de clave-valor son Riak y Berkeley DB. En algunos almacenes de clave-valor, como Redis, cada valor puede tener un tipo, como «entero», lo que le añade funcionalidad.

**Column-family:** Bases de datos orientadas a columnas, estas bases de datos, como Cassandra o HBase, permiten realizar consultas en grandes conjuntos de datos y almacenan los datos en columnas, en lugar de filas.

**Graph:** Bases de datos orientadas a grafo, se utilizan para almacenar información sobre redes de datos, como las conexiones sociales. Ejemplos de almacenes de grafos son Neo4J y Giraph.

**Document:** Bases de datos de documentos: en estas bases de datos se empareja cada clave con una estructura de datos compleja que se denomina 'documento'. Los documentos pueden contener muchos pares de clave-valor distintos, o pares de clave-matriz, o incluso documentos anidados.

Fuente: <https://www.mongodb.com/es/nosql-explained>.

**Glosario Técnico**

* **SQL:** Siglas provenientes del inglés: structured query language, que corresponden a la forma en la cual se denomina al lenguaje de programación SQL. El lenguaje de consulta estructurada hace referencia a la forma casi natural de la sintaxis SQL para trabajar con objetos y datos de una base de datos.
* **Base de datos relacional:**Elemento que almacena información respetando una estructura determinada y normalizada de tablas, campos y registros, usualmente relacionados entre sí, de manera tal que se mantenga la consistencia de la información allí alojada.
* **Base de datos NO-SQL:** Siglas provenientes del inglés: not only SQL, que hacen referencia a un tipo de base de datos no relacionales. Nacieron y se popularizaron fuertemente a partir de 2010, y conforman documentos al tipo de tablas donde se almacena la información. A diferencia de SQL, estas bases de datos pueden tener una estructura diferente dentro de un mismo documento, que difiere entre sus registros.
* **Modelo de Datos:**
* **Servidor de base de datos:**
* **Tablas:**Objeto de una base de datos donde se almacena la información, en forma de filas y columnas. Las filas son denominadas registros mientras que las columnas corresponden a los campos descriptivos de la información allí almacenada.
* **Registros:**Tipo de información que es almacenada en una tabla, respetando la estructura y tipos de datos definida por los campos.
* **SGBD:**Siglas correspondientes a Sistema de Gestión de Base de Datos. Hacen referencia a la aplicación de software que permite la gestión total de la información de las bases de datos. Usualmente estos sistemas son aplicaciones visuales que facilitan la interacción con los diferentes objetos de una DB.
* **DDL:**Siglas correspondientes a Data Definition Language. Hacen referencia al lenguaje SQL utilizado para crear y definir la estructura de los objetos de una base de datos: tablas, vistas, procedimientos almacenados, funciones, etc.
* **DML:**Siglas correspondientes a Data Manipulation Language. Hacen referencia a las sentencias SQL que nos permiten manipular la información almacenada en una base de datos (select, insert, update, delete).
* **DCL:**Siglas correspondientes a Data Control Language. Hacen referencia a los comandos SQL que permiten al administrador de una DB gestionar los datos almacenados en esta, a través de los comandos (grant, revoke) asociados a los usuarios de la DB.
* **TCL:**Siglas correspondientes a Transaction Control Language. Hacen referencia a los comandos que permiten confirmar o cancelar operaciones con datos en DB sql, a través de los comandos (begin transaction, commit, rollback).