Índice de contenido

[**¿Qué es un curador de datos?** 3](#_Toc128496664)

[**Bases de datos digitales** 3](#_Toc128496665)

[**Bases de datos**. 3](#_Toc128496666)

[**Integridad de datos**. 3](#_Toc128496667)

[**Ética**. 3](#_Toc128496668)

[**Tipos de bases de datos** 3](#_Toc128496669)

[**Bases de datos relacionales**. 3](#_Toc128496670)

[**Bases de datos estáticas**. 4](#_Toc128496671)

[**Bases de datos dinámicas**. 4](#_Toc128496672)

[**Estructura** 4](#_Toc128496673)

[**Modelo Entida-Relacion** 4](#_Toc128496674)

[**Normalización** 5](#_Toc128496675)

[Primera forma normal. 5](#_Toc128496676)

[Segunda forma normal. 5](#_Toc128496677)

[Tercera forma norma. 5](#_Toc128496678)

[(Tercera/Cuarta) forma normal. 6](#_Toc128496679)

[**Ejemplo 1 de normalización**. 6](#_Toc128496680)

[**Ejemplo 2 de normalización**. 7](#_Toc128496681)

[**Normalización** 7](#_Toc128496682)

[**Visualizador de datos** 8](#_Toc128496683)

[**Importar datos** 8](#_Toc128496684)

[**Exportar datos** 8](#_Toc128496685)

[**Depuración de datos** 8](#_Toc128496686)

[**Arreglo de datos** 9](#_Toc128496687)

[**Grafica dinámica** 9](#_Toc128496688)

[**Sistemas gestores** 9](#_Toc128496689)

[Sistemas gestores 9](#_Toc128496690)

[Adaptación al modelo de datos 9](#_Toc128496691)

[Arquitectura 10](#_Toc128496692)

[Componentes 10](#_Toc128496693)

[Apariencia de un sistema gestor 11](#_Toc128496694)

[**Usuarios y privilegios** 11](#_Toc128496695)

[Crear un usuario administrador 11](#_Toc128496696)

[Tipos de privilegios de los usuarios 11](#_Toc128496697)

[Crear usuarios y asignar o revocar privilegios 12](#_Toc128496698)

[**Lenguaje SQL** 12](#_Toc128496699)

[**Antes de crear una base de datos** 12](#_Toc128496700)

# **¿Qué es un curador de datos?**

Un curador de datos es la persona que se dedica a manipular, limpiary optimizarla información que se encuentra en una base de datos por medio de un software especializado. El curador de datos debe ser capaz de:

* *Crear bases* de datos para facilitar el almacenamiento y la consulta de grandes cantidades de información.
* *Limpiar la información* para que las bases funcionen más rápido y ocupen menos espacio.
* *Dar mantenimiento* a las bases mediante el respaldo y recuperación de la información.
* *Presentar al cliente informes* de resultados por medio de tablas y gráficas.

# **Bases de datos digitales**

## **Bases de datos**.

Las bases de datos han pasado de estar en papel a digitalizarse. Funcionan para guardar información, por ejemplo:

* Una tienda registra a los clientes y productos que se venden, de esta manera la tienda:
  + Sabe si llenar el almacén o no.
  + Que producto se vendió a que cliente.
  + Que factura pertenece a cierto producto.

## **Integridad de datos**.

* Proporciona una identificación única para cada registro.
* Al guardar la información: No se repite ni se pierde.
* Seguridad: Solo las personas autorizadas tienen acceso a la información.

## **Ética**.

En una base de datos se maneja información confidencial, por ende:

* No se debe compartir información sin autorización.
* No copiar ni robar información.
* Notificar si hay alguna anomalía.
* Todo se debe hacer con autorización.

# **Tipos de bases de datos**

## **Bases de datos relacionales**.

Son las bases de datos más utilizadas, sus características son:

* + Guarda los registros en tablas.
  + Pueden ser utilizadas por cualquier usuario.
  + Su gestión es fácil.
  + Se puede acceder rápidamente a los datos.
  + Garantiza la total consistencia de los datos, sin posibilidad de error.

## **Bases de datos estáticas**.

Estas bases de datos son de tipo lectura, permiten:

* + Almacenamiento de datos: Pueden guardar información..
  + Solo lectura: No permite modificar o eliminar la información que se almacena.
  + Combinación de datos: En ocasiones se suele hacer una combinación entre varias bases de datos estáticas, según las necesidades. Por ejemplo, podemos ver la evolución en el tiempo de distintas variables.

## **Bases de datos dinámicas**.

Las bases de datos dinámicas tienen flexibilidad, permite:

* + Almacenamiento de datos: Pueden guardar información.
  + Flexible: Permite modificar y eliminar la información que se almacena.
  + Necesidad de establecer relaciones: En este tipo de base de datos es importante establecer relaciones entre entidades.
  + Información cambiante: Se orienta a información que puede cambiar con el tiempo.

# **Estructura**

Las bases de datos se conforman por:

* + Tablas: Las tablas representan una entidad y contienen registros.
* Filas: Son cada uno de los registros de la tabla.
  + Llave primaria: Cada registro tiene un identificador único para evitar que los registros se repitan.
  + Campos/Columnas: Son las columnas de la tabla y representan las celdas de cada registro, (los atributos).

## **Modelo Entida-Relacion**

Entidad/Tabla: Representa cualquier cosa y es una tabla de la base de datos.

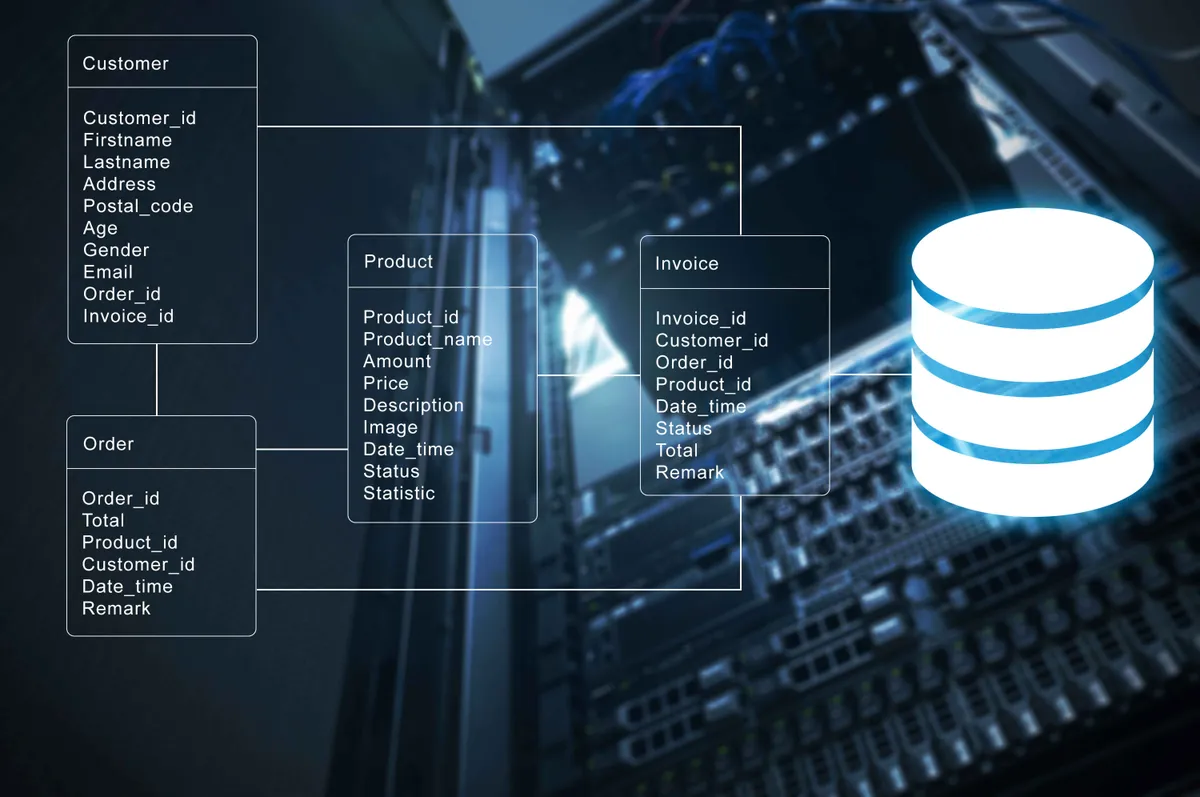
* Filas: Representan cada registro de la tabla.
  + Columnas: Representan cada atributo de los registros.

Las tablas pueden relacionarse entre sí por medio de su llave primaria, por ejemplo:

* Una venta se relaciona con un producto.
* Una categoría se relaciona con varios productos.
* Etc.

Las tablas que hay serían:

* Tabla: Categorías.
* Tabla: Productos.
* Tabla: Ventas.



# **Normalización**

Trata sobre organizar los datos por medio de su modelo Entidad-Relacion.

Existen varias reglas que deben seguirse para ordenar los datos:

## Primera forma normal.

* Eliminar los grupos repetidos de las tablas individuales.
* Crear una tabla independiente para cada conjunto de datos relacionados.
* Crear una llave primaria para cada registro, (clave única).
* Los campos de los registros no deben contener campos nulos.

## Segunda forma normal.

* Crear tablas individuales para guardar la información repetida de los campos de los registros.
* Relacionar la informacion de las tablas con los registros por medio de una llave foranea, (clave externa).

## Tercera forma norma.

* Borrar los campos que no dependan de la llave primaria.

## (Tercera/Cuarta) forma normal.

* La cuarta forma también llamada Forma normal de Boyce Codd (BCNF, Boyce Codd Normal Form) y la quinta también existen. Pero rara vez se consideran en un diseño real. Si no se aplican estas reglas, el diseño de la base de datos puede ser menos perfecto, pero no debería afectar a la funcionalidad.

## **Ejemplo 1 de normalización**.

1. **Tabla sin normalizar.**

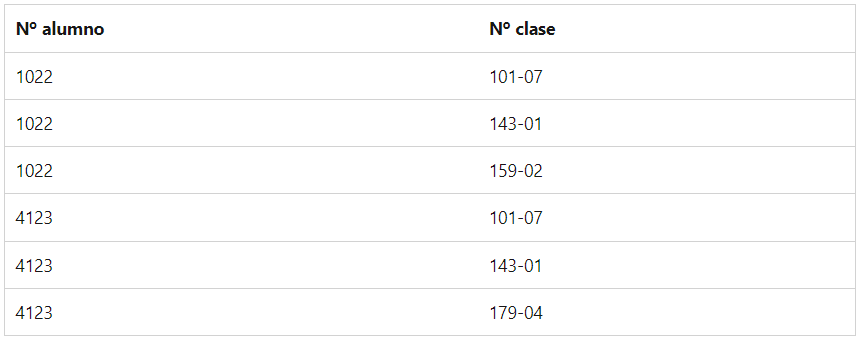
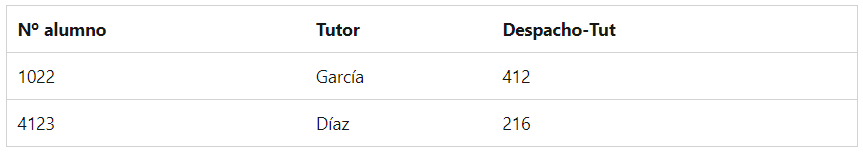


1. **Primera forma normal:** Sin grupos repetidos.

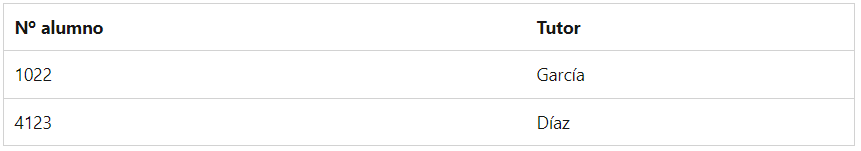
****

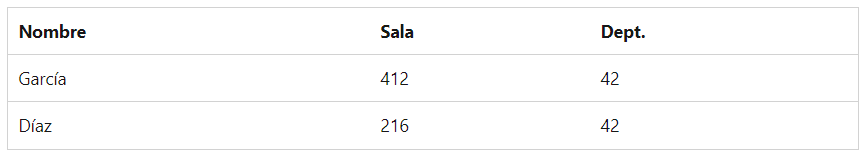
1. **Segunda forma normal:** Eliminar datos redundantes.

Tabla de alumno.

Tabla de registro.

**Tercera forma normal**: Eliminar los datos que no dependen de la clave.

Alumno.

Personal.

## **Ejemplo 2 de normalización**.

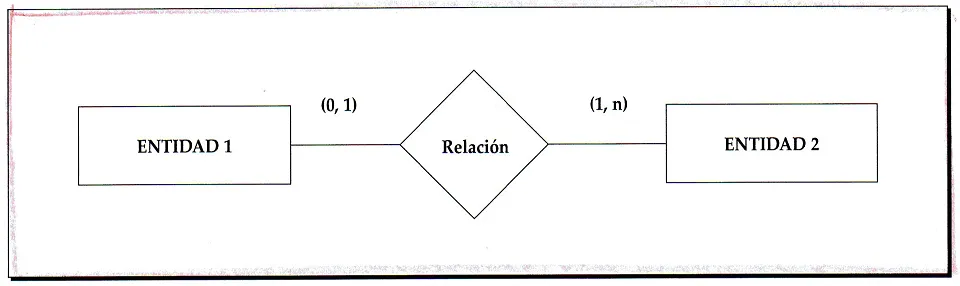
# **Normalización**

La cardinalidad es la forma en que se relacionan los registros de las tablas, los registros no pueden relacionarse con los registros de la misma tabla, deben relacionarse con un registro de una tabla diferente:

1. (**Uno a uno, [1:1]**), un registro solo puede relacionarse con uno.
2. (**Uno a muchos, [1:N]**), un registro puede relacionarse con muchos.
3. (**Muchos a muchos, [N:M]**), dos o más registros se pueden relacionar con dos o más registros.

Por ejemplo:

1. [**1:1**]. Un perfil pertenece solo a un usuario y un usuario solo puede tener un perfil.
2. [**1:N**]. Una categoría tiene muchos productos y muchos productos pueden estar en una categoría.
3. [**N:M**]. Una venta puede tener muchos productos, un producto puede estar en varias ventas.



En la imagen anterior podemos ver que una entidad puede relacionarse con otra haciendo uso de la cardinalidad.

# **Visualizador de datos**

## **Importar datos**

Datos > Desde texto > “Buscas tu archivo txt” > Importar

> Opción delimitados > Siguiente > Seleccionar solo el separados coma.

> Calificador de texto en “comillas dobles”. > Siguiente.

> Seleccionamos “general” > Finalizar.

Ahora solo colocamos la celda en donde aparecerá la tabla, ejemplo:

*=$A$1* igual a *A1*. > Aceptar.

## **Exportar datos**

Guardar como > CSV delimitado por comas > Aceptar.

## **Depuración de datos**

*Eliminar valores duplicados.*

Seleccionar cualquier celda de la tabla > Insertar > Tabla

> Seleccionar que la tabla tiene encabezados > Aceptar.

*Se creara una tabla que podemos depurar.*

Seleccionamos la tabla > Formato condicional > Resaltar reglas de celdas

> Valores duplicados > Elegimos el color de los valores duplicados

> Aceptar.

Datos > Quitar duplicados

> Seleccionar campos que se desean comparar > Aceptar.

## **Arreglo de datos**

*Tabla dinámica*

Seleccionar cualquier celda previamente depurada > Insertar > Tabla

Dinámica > Confirmar que todas las celdas estén seleccionadas

> Opción nueva hoja de cálculo > Aceptar.

Elegir el tipo de acomodo, por ejemplo:

* Filtro > Fecha.
* Columna > Precio.
* Fila > Producto.
* Valores > Podemos mostrar el valor total que queramos.

## **Grafica dinámica**

Seleccionar una celda de la tabla dinámica > Insertar > Grafica dinámica

> Se pueden modificar los parámetros: Filtros, Valores, Ejes, Leyendas.

Para guardar cambios de las modificaciones ya sea en la tabla dinámica o grafica dinámica, se debe:

Modificar la tabla originar > Actualizar todo.

# **Sistemas gestores**

## Sistemas gestores

Los sistemas gestores facilitan el uso de la información, permiten:

* Crear bases de datos.
* Manipula su contenido.
* Da seguridad a la información.
* Controla quien puede ver o manipular la información.
* Recupera y restaura los datos en caso de alguna falla.

## Adaptación al modelo de datos

Existen varias formas de adaptar los datos:

* Jerárquico.
* Red.
* Relacional.

Por número de usuarios conectados:

* Monousuario: Permite que solo un usuario esté conectado a la base de datos.
* Multiusuario: Dos o más usuarios pueden estar conectados al mismo tiempo.
* Centralizados: Todas las bases de datos están en una misma computadora.
* Distribuidos: Solo un conjunto de base de datos que pueden comunicarse entre sí.

## Arquitectura

Un gestor de base de datos puede dividirse en tres niveles:

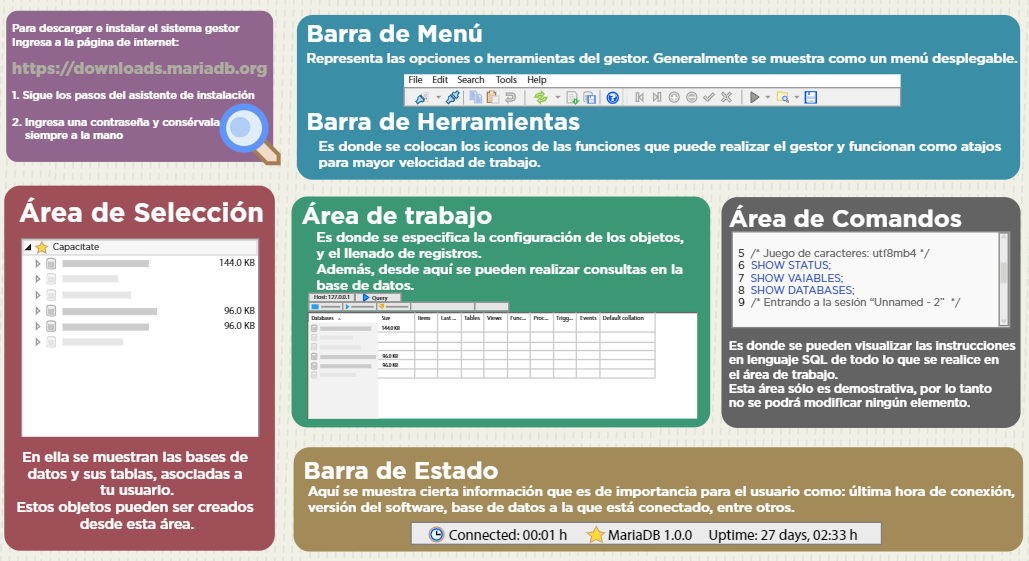
* Físico: El dispositivo de almacenamiento, guarda la información como un conjunto de bytes.
* Lógico: El lenguaje de programación que se utiliza para acceder al nivel físico.
* Externo: Es la interfaz gráfica para el usuario.

## Componentes

Los sistemas gestores tienen varias características:

* Usuarios:
  + Administradores: Deciden qué información habrá y dan apoyo técnico.
  + Programadores: Utilizan los lenguajes para manipular las bases de datos.
  + Usuario final: Es el que introduce, solicita y actualiza la información.
* Lenguajes: El conjunto de instrucciones que utilizan los administradores para manipular las bases de datos.
* Diccionario de datos: Guarda las definiciones de la información almacenada, (tipos de datos que se almacenan).
* Elementos de seguridad: Proporciona los mecanismos para garantizar la seguridad e integridad de los datos.

## Apariencia de un sistema gestor



# **Usuarios y privilegios**

## Crear un usuario administrador

Para crear un usuario administrado dentro del gestor de la base de datos, es necesario hacer lo siguiente:

1. Crear nuestro usuario como ROOT o algún otro nombre en el gestor.
2. Establecer una contraseña.

## Tipos de privilegios de los usuarios

Además del usuario administrador pueden existir otros usuario más, estos usuarios pueden tener privilegios que sólo el administrador puede asignar. *Los privilegios son todo lo que un usuario tiene permitido hacer para manipular las bases de datos*.

El usuario administrador por defecto tiene el privilegio global, que es el privilegio más alto y establece todos los privilegios que se muestran a continuación:

1. Privilegios de selección (SELECT): Permite al usuario leer los datos de una tabla o vista.
2. Privilegios de inserción (INSERT): Permite al usuario insertar nuevos datos en una tabla.
3. Privilegios de actualización (UPDATE): Permite al usuario modificar los datos existentes en una tabla.
4. Privilegios de eliminación (DELETE): Permite al usuario eliminar datos de una tabla.
5. Privilegios de ejecución (EXECUTE): Permite al usuario ejecutar procedimientos almacenados y funciones definidos en la base de datos.
6. Privilegios de creación (CREATE): Permite al usuario crear nuevos objetos en la base de datos, como tablas, vistas, índices, etc.
7. Privilegios de modificación (ALTER): Permite al usuario modificar la estructura de un objeto existente, como una tabla o vista.
8. Privilegios de eliminación (DROP): Permite al usuario eliminar un objeto existente de la base de datos.
9. Privilegios de concesión de privilegios (GRANT): Permite al usuario conceder privilegios a otros usuarios en la base de datos.
10. Privilegios de revocación de privilegios (REVOKE): Permite al usuario revocar los privilegios que ha otorgado previamente a otros usuarios.

## Crear usuarios y asignar o revocar privilegios

Podemos darle privilegios a uno o más usuarios que creemos dentro del gestor para manipular las bases de datos, debemos seguir los siguientes pasos:

1. Establecer el nombre del usuario a crear.
2. Establecer la contraseña del usuario a crear.
3. Establecer los privilegios de manipulación del usuario para las bases de datos.

También pueden revocarse los privilegios de un usuario siguiendo las instrucciones:

1. Revocar los privilegios de unos usuarios o darle más privilegios.
2. Eliminar al usuario.

# **Lenguaje SQL**

El lenguaje estructurado o de consultas SQL, Structured Query Language se divide en dos partes:

* Lenguaje de definición: Sirve para crear, modificar, borrar tablas, campos o una base de datos.
* Lenguaje de manipulación: Sirve para crear, modificar, buscar y borrar los registros.

# **Antes de crear una base de datos**

Antes de comenzar a crear las tablas, es necesario tener identificados los siguientes puntos:

* Las entidades que quieres crear.
* Los atributos de las entidades.
* Las relaciones que habrá entre cada entidad.