

Taller 1
Bases de datos
Universidad de Medellín
Víctor Pérez Cano 1036681583

1. Realizar un cuadro comparativo donde presente las ventajas y desventajas de las bases de datos relacionales.

Ventaja	Desventaja
Provee herramientas que garantizan evitar la duplicidad de registros.	Presentan deficiencias con datos gráficos, multimedia, CAD y sistemas de información geográfica.
Garantiza la integridad referencial, así, al eliminar un registro elimina todos los registros relacionados dependientes.	No se manipulan de forma manejable los bloques de texto como tipo de dato.
Favorece la normalización por ser más comprensible y aplicable	Las bases de datos orientadas a objetos (BDOO) se propusieron con el objetivo de satisfacer las necesidades de las aplicaciones anteriores y así, complementar, pero no sustituir a las bases de datos relacionales.
Mejora en la seguridad y en la integridad de datos	Costosa creación y mantenimiento del sistema

2. Explicar en qué consisten las 3 formas normales del modelo relacional en bases de datos. Utilizar ejemplos.

La Primera Forma Normal: Esta primera Forma Normal, nos lleva a no repetir datos en nuestras tablas. Los famosos maestro – detalle, deben aplicarse a la estructura de la tabla. Si nuestra tabla de ventas repite una y otra vez (por cada venta), el nombre, el domicilio y otros datos del Cliente, es que no hemos aplicado esta Normalización.

Ejemplo:

Si tenemos una tabla clientes, en la tabla ventas, solo debería figurar el código del cliente, para que el resto de los datos se puedan referenciar automáticamente sin problemas y sin duplicar información. Lo mismo ocurriría en una tabla de detalle de ventas, si por cada ítem vendido colocamos el detalle del producto, con su descripción, medidas, etc. Tendríamos un desaprovechamiento de espacio y recursos muy grande. Para ello, tendremos nuestra tabla maestra de Productos y con solo grabar el código de dicho producto en nuestra tabla de ventas, será suficiente.

La Segunda Forma Normal: (Si o si debe estar previamente aplicada la Primera Forma Normal) La Segunda Forma Normal nos habla de que cada columna de la tabla debe depender de la clave. Esto significa que todo un registro debe depender únicamente de la clave principal, si tuviéramos alguna columna que se repite a lo largo de todos los registros, dichos datos deberían dividirse en una nueva tabla.

Ejemplo:

VentaID	ItemID	FechaVenta	ClienteVenta	Productoid	Cantidad
1	1	01/12/2007	2	2334	10
1	2	01/12/2007	2	3333	2
1	3	01/12/2007	2	66643	34
1	4	01/12/2007	2	21	3
2	1	02/12/2007	5	3566	6

En el modelo anterior se presenta un problema, se busca no repetir datos. Si toda venta tendrá el mismo número de cliente y la misma fecha se puede crear una tabla y que contenga esos 2 datos. Es evidente que la columna **ClienteVenta** y **FechaVenta** se repetirán por cada venta realizada. Es por ello que se propone el siguiente esquema:

VentaID	ItemID	Productoid	Cantidad
1	1	2334	10
1	2	3333	2
1	3	66643	34
1	4	21	3
2	1	3566	6

Y ahora nuestra nueva tabla maestra:

VentaID	FechaVenta	ClienteVenta
1	01/12/2007	2
2	02/12/2007	5

Entonces, nuestra 2da Forma Normal nos habla de que cada columna de una tabla debe depender de toda la clave y no constituir un dato único para cada grupo de registros.

La Tercera Forma Normal: En realidad si nos guiamos en el ejemplo de esta nota, ya no quedaría normalización por aplicar y podríamos decir que nuestro ejemplo cumple con las 3 formas normales, ya que la 3ra Forma Normal nos habla de que:

1. Ninguna Columna puede depender de una columna que no tenga una clave
2. No puede haber datos derivados

En el 2do ejemplo hemos descubierto campos que dependían de la clave principal (VentaID) y que podrían incluirse en una tabla maestra. Pero supongamos un ejemplo donde ciertas columnas no dependen de la clave principal y si dependen de una columna de nuestra tabla.

VentaID	ItemID	ProductoID	Cantidad	Descripcion	Medida	Proveedor
1	1	3455	12	Impresora HP LJ8000	122cm	1
1	2	2455	34	Scanner HP A3555	33cm	1
2	1	5444	21	Mouse HP Wireless	–	1

Esto es muy normal encontrar en bases mal normalizadas. Vemos que los campos DESCRIPCION, MEDIDA y PROVEEDOR no dependen de VENTAID y es por ello que no deberían estar dentro de la tabla de detalle de ventas, ya que dependen de PRODUCTOID. Aquí no se trata ya de eliminar grupos repetidos de datos (1ra Forma Normal) sino que, ante la inclusión de una clave perteneciente a otra tabla, cualquier campo que sea subordinado de dicha clave debe estar en otra tabla y no en nuestra tabla detalle.

3. Enumerar y explicar 10 ventajas de los SGBD.

- **Aumento de la concurrencia:** La mayoría de los SGBD gestionan el acceso concurrente a la base de datos y garantizan que no ocurran problemas en el acceso de múltiples usuarios.
- **Mejora en la accesibilidad a los datos** Muchos SGBD proporcionan lenguajes de consultas o generadores de informes que permiten al usuario hacer cualquier tipo de consulta sobre los datos, sin que sea necesario que un programador escriba una aplicación que realice tal tarea.
- **Posibilidad de compartir datos:** Eliminando o controlando las redundancias de datos se reduce en gran medida el riesgo de que haya inconsistencias. Si un dato está almacenado una sola vez, cualquier actualización se debe realizar sólo una vez, y está disponible para todos los usuarios inmediatamente. Si un dato está duplicado y el sistema conoce esta redundancia, el propio sistema puede encargarse de garantizar que todas las copias se mantienen consistentes. Desgraciadamente, no todos los SGBD de hoy en día se encargan de mantener automáticamente la consistencia.
- **Mejor rendimiento:** Las bases de datos de cada uno de los sitios son más pequeñas. Las consultas locales y las transacciones que tienen acceso a un solo sitio son más rápidas. Las transacciones que acceden a mas de un sitio se ejecutan en paralelo.
- **Consistencia de datos:** Eliminando o controlando las redundancias de datos se reduce en gran medida el riesgo de que haya inconsistencias. Si un dato está almacenado una sola vez, cualquier actualización se debe realizar sólo una vez, y está disponible para todos los usuarios inmediatamente. Si un dato está duplicado y el sistema conoce esta redundancia, el propio sistema puede encargarse de garantizar que todas las copias se mantienen consistentes. Desgraciadamente, no todos los SGBD de hoy en día se encargan de mantener automáticamente la consistencia.
- **Control sobre la redundancia de datos:** Los sistemas de archivos almacenan varias copias de los mismos datos en archivos distintos. Esto hace que se desperdicie espacio de almacenamiento, además de provocar la falta de consistencia de datos. En los sistemas de bases de datos todos estos archivos están integrados, por lo que no se almacenan varias copias de los mismos datos. Sin embargo, en una base de datos no se puede eliminar la redundancia completamente, ya que en ocasiones es necesaria para modelar las relaciones entre los datos, o bien es necesaria para mejorar las prestaciones.
- **Mantenimiento de estándares:** Gracias a la integración es más fácil respetar los estándares necesarios, tanto los establecidos a nivel de la empresa como los nacionales e internacionales.

Estos estándares pueden establecerse sobre el formato de los datos para facilitar su intercambio, pueden ser estándares de documentación, procedimientos de actualización y también reglas de acceso.

- **Mejora en la accesibilidad a los datos:** Muchos SGBD proporcionan lenguajes de consultas o generadores de informes que permiten al usuario hacer cualquier tipo de consulta sobre los datos, sin que sea necesario que un programador escriba una aplicación que realice tal tarea.
- **Más información sobre la misma cantidad de datos:** Al estar todos los datos integrados, se puede extraer información adicional sobre los mismos.
- **Mejora en la seguridad:** Los SGBD permiten mantener la seguridad mediante el establecimiento de claves para identificar al personal autorizado a utilizar la base de datos. Las autorizaciones se pueden realizar a nivel de operaciones, de modo que un usuario puede estar autorizado a consultar ciertos datos, pero no a actualizarlos.

4. ¿Qué es una transacción en un SGBD? ¿Qué garantías ofrecen los SGBD con respecto a las transacciones?

Transacción:

Es un conjunto de órdenes que se ejecutan formando una unidad de trabajo, es decir, en forma indivisible o atómica. Que es capaz de mantener la integridad de los datos.

Garantías:

Es capaz de mantener la integridad de los datos, haciendo que estas transacciones no puedan finalizar en un estado intermedio

5. ¿Cuándo se deben guardar los datos en SGBD en lugar de hacerlo en los archivos del sistema operativo y viceversa?

Se deben guardar los datos en SGBD cuando se desee ver la información mediante múltiples usuarios y así poder acceder a ellas al mismo tiempo ya que es más eficiente que usar archivos del sistema operativo.

6. ¿Qué son los modelos de datos? ¿Qué es el modelo de datos relacional? ¿Qué es la independencia con respecto a los datos y cómo la soportan los SGBD?

El modelo de datos relacional es aquel en que todos los datos son almacenados en relaciones, y como cada relación es un conjunto de datos, el orden en el que éstos se almacenen no tiene relevancia, sirve para modelar problemas reales y administrar datos dinámicamente y utiliza números sistemas como Oracle, Paradox, etc.

7. Explicar las ventajas de emplear un lenguaje de consultas para procesar los datos en lugar de los programas habituales.

- Reducción en tiempo de desarrollo.
- Buen mantenimiento
- Eficacia y rapidez

8. ¿Qué es la independencia lógica con respecto a los datos y por qué es importante?

Es la capacidad de modificar el esquema conceptual sin tener que alterar los esquemas externos ni los programas de aplicación. Podemos modificar el esquema conceptual para ampliar la base de datos o para reducir la base de datos.

Es importante porque en un SGBD permite que los programas queden aislados de las modificaciones debido al modo en que se estructuran y se guardan los datos.

9. Explicar la diferencia entre la independencia lógica con respecto a los datos y la física.

En la independencia lógica la estructura lógica si se modifica un esquema, las aplicaciones no cambian en cambio la independencia física es la capacidad de modificar el esquema interno sin necesidad de alterar el sistema conceptual.

10. Explicar las diferencias entre los esquemas externo, interno y conceptual. ¿Cómo están relacionadas estas capas de esquemas con los conceptos de independencia lógica y física con respecto a los datos?

La diferencia entre los diferentes esquemas es que el esquema externo es el que se encarga de personalizar y dar el acceso a los datos del usuario o cualquier grupo semejante de este, así como dar una o varias vistas de la relación, mientras tanto el esquema interno o físico indica en resumen el modo en que las relaciones se guardan en dispositivos de almacenamiento; por lo tanto el esquema conceptual es aquel que describe todas las relaciones almacenadas en las base de datos.

El esquema externo se relaciona con la independencia lógica gracias a que el usuario es el encargado de pedir la relación de vistas de acuerdo a las relaciones almacenadas en las bases de datos del esquema conceptual.

11. ¿Cuáles son las responsabilidades de los DBA?

- 1) Instalación y actualización del software manejador de base de datos.
- 2) Asignación de recursos para la utilización de la base de datos: memoria, espacio en disco, perfiles de usuario etc.
- 3) Ajuste de la base de datos para conseguir el rendimiento óptimo.

4) Estrategias de copia de seguridad y recuperación.

5) Colaboración con el personal de administración del sistema y desarrolladores de aplicaciones.

12. ¿Cuál de los elementos siguientes desempeña un papel importante en la representación de la información sobre el mundo real en las bases de datos? Explicar brevemente.

- a) El lenguaje de definición de datos.
- b) El lenguaje de manipulación de datos.
- c) El gestor de la memoria intermedia.
- d) El modelo de datos.

d) El modelo de datos es el que determina la estructura lógica de una base de datos y de manera fundamental determina el modo de almacenar, organizar y manipular los datos. Su construcción se basa en la manera en la que está constituida la información en el mundo real y cómo se comporta dentro del contexto.

13. ¿Por qué se utiliza el modelo ER para crear el diseño inicial?

El modelo entidad relación es utilizado en el diseño inicial ya que ayuda a identificar los objetos del mundo real que deben ser entidades dentro del modelo y diseñar una forma más correcta y eficaz con la que se deben relacionar las diferentes entidades, así como los atributos que deben contener cada una de las entidades para resolver el problema de base de datos.

14. ¿Cuáles son los conceptos principales del modelo ER?

Entidad. La entidad es cualquier clase de objeto o conjunto de elementos presentes o no, en un contexto determinado. Dicho de otra forma, las entidades las constituyen las tablas de la base de datos que permiten el almacenamiento de los ejemplares o registros del sistema, quedando recogidos bajo la denominación o título de la tabla o entidad.

Atributos. Son las características, rasgos y propiedades de una entidad, que toman como valor una instancia particular. Es decir, los atributos de una tabla son en realidad sus campos descriptivos.

Relación. Vínculo que permite definir una dependencia entre los conjuntos de dos o más entidades. Esto es la relación entre la información contenida en los registros de varias tablas.

Interrelación. Las interrelaciones las constituyen los vínculos entre entidades, de forma tal que representan las relaciones definidas en el esquema relacional de forma efectiva. Esto no sólo la relación de los registros sino de sus tablas y de las características de la interrelación entre las entidades, a través de un campo clave que actúa como código de identificación y referencia para relacionar (es decir, como nexo de unión y articulación de la relación). Los tipos de interrelaciones entre entidades o tablas se realizan aplicando las reglas de cardinalidad y modalidad.

Entidades fuertes. Lo constituyen las tablas principales de la base de datos que contienen los registros principales del sistema de información y que requieren de entidades o tablas auxiliares para completar su descripción o información. Por ejemplo, la tabla usuario es una entidad fuerte en relación a la tabla tipos de usuarios, que es una entidad débil dada su condición auxiliar para clasificar a los usuarios registrados en la biblioteca.

Entidades débiles. Son entidades débiles a las tablas auxiliares de una tabla principal a la que completan o complementan con la información de sus registros relacionados. Por ejemplo, también son consideradas entidades débiles las tablas intermedias que sirven para compartir información de varias tablas principales.

Integridad referencial. Se denomina integridad referencial al tipo de interrelación que se produce entre tablas mediante un campo clave que deberá contener la cadena alfanumérica exacta al identificador de la tabla auxiliar para poder realizar la relación entre los registros. En caso contrario no se produce la relación. Además, se trata de un mecanismo que evita duplicidades e incorrecciones ya que la propiedad de integridad referencial conmina a que los datos de un usuario además de su identificador ID sean distintos al de los demás. Dicho de otra forma, no pueden existir dos registros iguales con los mismos datos.

Tipos de relaciones:

- **Relación 1 a 1.** La relación uno a uno, define que un único registro de la tabla puede estar relacionado con un único registro de la tabla relacionada.
- **Relación 1 a *.** La relación de uno a varios, define que un registro dado de una tabla auxiliar o secundaria sólo puede estar vinculado con un único registro de la tabla principal con la que está relacionada.
- **Relación * a *.** La relación de varios a varios, define que un registro de una tabla puede estar relacionado con varios registros de la tabla relacionada y viceversa.

Modelo entidad relación	Objeto de la base de datos	Ejemplo
Entidad	Tabla de la base de datos	Tabla usuarios
Atributos – Intención	Campos de una tabla	id, nombre, apellidos, tipo de usuario, DNI, dirección, teléfono
Relación	Vínculo entre conjuntos	Jorge Martínez es investigador
Interrelación	Relación entre tablas	<i>Tabla Usuarios</i> relacionada con <i>Tabla Tipo de usuarios</i>

Entidades fuertes	Tabla principal	Tabla Usuarios
Entidades débiles	Tabla auxiliar	Tabla Tipo de usuarios

Esquema con algunos elementos fundamentales del diagrama ER

15. Explicar brevemente los términos siguientes: atributo, dominio, entidad, relación, conjunto de entidades, conjunto de relaciones, relación de una a varias, relación de varias a varias, restricción de participación, restricción de solapamiento, restricción de cobertura, conjunto de entidades débiles, agregación e indicador de papel.

- **Atributo:** Son las características, rasgos y propiedades de una entidad, que toman como valor una instancia particular. Es decir, los atributos de una tabla son en realidad sus campos descriptivos.
- **Dominio:** Es el conjunto de valores permitidos para cada atributo.
- **Entidad:** La entidad es cualquier clase de objeto o conjunto de elementos presentes o no, en un contexto determinado. Dicho de otra forma, las entidades las constituyen las tablas de la base de datos que permiten el almacenamiento de los ejemplares o registros del sistema, quedando recogidos bajo la denominación o título de la tabla o entidad.
- **Relación:** Vínculo que permite definir una dependencia entre los conjuntos de dos o más entidades. Esto es la relación entre la información contenida en los registros de varias tablas.
- **Conjunto de entidades:** Un conjunto de entidades es un contenedor lógico para instancias de un tipo de entidad y las instancias de cualquier tipo derivado de ese tipo de entidad. La relación entre un tipo de entidad y un conjunto de entidades es análoga a la relación entre una fila y una tabla en una base de datos relacional: al igual que una fila, un tipo de entidad describe la estructura de los datos y, al igual que una tabla, un conjunto de entidades contiene instancias de una determinada estructura.
- **Conjunto de relaciones:** Conjunto de relaciones es un conjunto de relaciones del mismo tipo, por ejemplo, entre ARTÍCULOS y VENTAS todas las asociaciones existentes entre los artículos y las ventas que tengan estos, forman un conjunto de relaciones.
- **Relación de una a varias:** La relación de uno a varios, define que un registro dado de una tabla auxiliar o secundaria sólo puede estar vinculado con un único registro de la tabla principal con la que está relacionada.
- **Relación de varias a varias:** La relación de varios a varios, define que un registro de una tabla puede estar relacionado con varios registros de la tabla relacionada y viceversa.
- **Restricción de participación:**
 - **Participación total:** Cada entidad de un conjunto de entidades participa al menos en una relación del conjunto de relaciones.
 - **Participación parcial:** Sólo algunas entidades del conjunto de entidades participan en relaciones en el conjunto de relaciones.
- **Conjunto de entidades débiles:** Son entidades débiles a las tablas auxiliares de una tabla principal a la que completan o complementan con la información de sus registros relacionados. Por ejemplo, también son consideradas entidades débiles las tablas intermedias que sirven para compartir información de varias tablas principales.