

Base de datos relacional: Normalización de datos



La normalización de bases de datos es el proceso de estructurar una base de datos de manera que se reduzca la redundancia y se asegure la integridad de los datos. Las tres formas principales de normalización, también conocidas como las tres primeras formas normales (1NF, 2NF y 3NF), son esenciales para diseñar tablas relacionales eficientes y coherentes.

Primera Forma Normal (1NF)

Una tabla está en la Primera Forma Normal si cumple con las siguientes condiciones:

1. **Dominio de atributo atómico:** Todos los atributos contienen valores indivisibles. Es decir, cada columna debe contener valores atómicos, sin listas o conjuntos.
2. **Valores únicos por columna:** Cada columna debe contener valores únicos para cada fila en la tabla.
3. **Ausencia de duplicados:** No puede haber filas duplicadas en la tabla.

Dato Atómico

Un dato atómico es un valor que no puede ser dividido o descompuesto en partes más pequeñas dentro del contexto de una base de datos relacional. La atomicidad asegura que cada celda en una tabla contenga una única pieza de información. Esto es fundamental para la Primera Forma Normal (1NF).

Ejemplo de 1NF

Supongamos una tabla de clientes con direcciones múltiples no normalizada:

id	nombre	direcciones
1	Juan	Calle 1, Calle 2
2	María	Calle 3

Para convertirla a 1NF, necesitamos que cada dirección esté en una fila separada:

id	nombre	direccion
1	Juan	Calle 1
1	Juan	Calle 2
2	María	Calle 3

Segunda Forma Normal (2NF)

Una tabla está en la Segunda Forma Normal si cumple con las siguientes condiciones:

1. Está en 1NF.
2. Eliminación de dependencias parciales: Todos los atributos no clave deben depender completamente de la clave primaria. Es decir, no debe haber atributos que dependan solo de una parte de una clave primaria compuesta.

Ejemplo de 2NF

Supongamos una tabla de **pedidos** no normalizada:

id	producto_id	cantidad	nombre_producto
1	101	5	Producto A
1	102	3	Producto B
2	101	2	Producto A

En este caso, **nombre_producto** depende solo de **producto_id** y no de **id**. Para normalizarla a 2NF, separamos los datos en dos tablas:

tabla de pedidos

id	producto_id	cantidad
1	101	5
1	102	3
2	101	2

tabla de productos

id	nombre_producto
101	Producto A
102	Producto B

Tercera Forma Normal (3NF)

Una tabla está en la Tercera Forma Normal si cumple con las siguientes condiciones:

1. Está en 2NF.
2. Eliminación de dependencias transitivas: Todos los atributos no clave deben depender únicamente de la clave primaria y no de otros atributos no clave.

Ejemplo de 3NF

Supongamos una tabla de **empleados** no normalizada:

id	nombre	departamento_id	nombre_departamento
1	Ana	10	Ventas
2	Luis	20	Marketing
3	Marta	10	Ventas

En este caso, **nombre_departamento** depende de **departamento_id**, no directamente de **empleado_id**. Para normalizarla a 3NF, separamos los datos en dos tablas:

tabla de empleados

id	nombre	departamento_id
1	Ana	10
2	Luis	20
3	Marta	10

tabla de departamentos

id	nombre_departamento
10	Ventas
20	Marketing

En resumen, la normalización en 1NF elimina grupos repetitivos, en 2NF elimina dependencias parciales, y en 3NF elimina dependencias transitivas, lo que ayuda a mantener la integridad de los datos y reducir la redundancia en las bases de datos relacionales.