

- NORMALIZACION DE BASE DE DATOS -



La normalización de bases de datos es un proceso que consiste en designar y aplicar una serie de reglas a las relaciones obtenidas tras el paso del modelo entidad-relación al modelo relacional.

Las bases de datos relacionales se normalizan para:

- Evitar la redundancia de los datos.
- Disminuir problemas de actualización de los datos en las tablas.
- Proteger la integridad de datos.

En el modelo relacional es frecuente llamar tabla a una relación, aunque para que una tabla sea considerada como una relación tiene que cumplir con algunas restricciones:

- Cada tabla debe tener su nombre único.
- No puede haber dos filas iguales. No se permiten los duplicados.
- Todos los datos en una columna deben ser del mismo tipo.

I. FORMA NORMAL 1

Según la definición de Date de la 1FN, una tabla está en 1FN si y solo si es "isomorfa a alguna relación", lo que significa, específicamente, que satisface las siguientes cinco condiciones:

- No hay orden de arriba-a-abajo en las filas.
- No hay orden de izquierda-a-derecha en las columnas.
- No hay filas duplicadas.
- Cada intersección de fila-y-columna contiene exactamente un valor del dominio aplicable (y nada más).
- Todas las columnas son regulares [es decir, las filas no tienen componentes como IDs de fila, IDs de objeto, o timestamps ocultos].

La violación de cualquiera de estas condiciones significaría que la tabla no es estrictamente relacional, y por lo tanto no está en 1FN.

Algunos ejemplos de tablas (o de vistas) que no satisfacen esta definición de primera forma normal son:

- Una tabla que carece de una clave primaria. Esta tabla podría acomodar filas duplicadas, en violación de la condición 3.
- Una vista cuya definición exige que los resultados sean retornados en un orden particular, de modo que el orden de la fila sea un aspecto intrínseco y significativo de la vista. Esto viola la condición 1. Las tuplas en relaciones verdaderas no están ordenadas una con respecto de la otra.
- Una tabla con por lo menos un atributo que pueda ser nulo. Un atributo que pueda ser nulo estaría en violación de la condición 4, que requiere a cada campo contener exactamente un valor de su dominio de columna. Sin embargo, debe observarse que este aspecto de la condición 4 es controvertido. Muchos autores consideran que una tabla está en 1FN si ninguna clave candidata puede contener valores nulos, pero se aceptan éstos para atributos (campos) que no sean clave, según el modelo original de Codd sobre el modelo relacional, el cual hizo disposición explícita para los nulos.

II. FORMA NORMAL 2

La segunda forma normal (2NF) es una forma normal usada en normalización de bases de datos. La 2NF fue definida originalmente por E.F. Codd en 1971. Una tabla que está en la primera forma normal (1NF) debe satisfacer criterios adicionales para calificar para la segunda forma normal. Específicamente: una tabla 1NF está en 2NF si y solo si, dada una clave primaria y cualquier atributo que no sea un constituyente de la clave primaria, el atributo no clave depende de toda la clave primaria en vez de solo de una parte de ella.

En términos levemente más formales: una tabla 1NF está en 2NF si y solo si ninguno de sus atributos no-principales son funcionalmente dependientes en una parte (subconjunto propio) de una clave candidata (Un atributo no-principal es uno que no pertenece a ninguna clave candidata).

Observa que cuando una tabla 1NF no tiene ninguna clave candidata compuesta (claves candidatas consisten en más de un atributo), la tabla está automáticamente en 2NF.

III. FORMA NORMAL 3

La tercera forma normal (3NF) es una forma normal usada en la normalización de bases de datos. La 3NF fue definida originalmente por E.F. Codd en 1971. La definición de Codd indica que una tabla está en 3NF si y solo si las tres condiciones siguientes se cumplen:

La tabla está en la segunda forma normal (2NF)

Ningún atributo no-primario de la tabla es dependiente transitivamente de una clave primaria

Es una relación que no incluye ningún atributo clave

Un atributo no-primario es un atributo que no pertenece a ninguna clave candidata. Una dependencia transitiva es una dependencia funcional $X \rightarrow Z$ en la cual Z no es inmediatamente dependiente de X, pero sí de un tercer conjunto de atributos Y, que a su vez depende de X. Es decir, $X \rightarrow Z$ por virtud de $X \rightarrow Y$ e $Y \rightarrow Z$.

Una formulación alternativa de la definición de Codd, dada por Carlo Zaniolo en 1982, es ésta: Una tabla está en 3NF si y solo si, para cada una de sus dependencias funcionales $X \rightarrow A$, por lo menos una de las condiciones siguientes se mantiene:

X contiene A, ó

X es una superclave, ó

A es un atributo primario (es decir, A está contenido dentro de una clave candidata)

La definición de Zaniolo tiene la ventaja de dar un claro sentido de la diferencia entre la 3NF y la más rigurosa forma normal de Boyce-Codd (BCNF). La BCNF simplemente elimina la tercera alternativa ("A es un atributo primario").

IV. FORMA NORMAL 4

La cuarta forma normal (4NF) es una forma normal usada en la normalización de bases de datos. La 4NF se asegura de que las dependencias multivaluadas independientes estén correctas y eficientemente representadas en un diseño de base de datos. La 4NF es el siguiente nivel de normalización después de la forma normal de Boyce-Codd (BCNF).

Una tabla está en 4NF si y solo si está en Tercera forma normal o en BCNF (Cualquiera de ambas) y no posee dependencias multivaluadas no triviales. La definición de la 4NF confía en la noción de una dependencia multivaluada. Una tabla con una dependencia multivaluada es una donde la existencia de dos o más relaciones independientes muchos a muchos causa redundancia; y es esta redundancia la que es suprimida por la cuarta forma normal.

V. FORMA NORMAL 5

La quinta forma normal (5FN), también conocida como forma normal de proyección-uni3n (PJ/NF), es un nivel de normalizaci3n de bases de datos dise1ado para reducir redundancia en las bases de datos relacionales que guardan hechos multi-valores aislando sem1nticamente relaciones m1ltiples relacionadas. Una tabla se dice que est1 en 5NF si y s3lo si est1 en 4NF y cada dependencia de uni3n (join) en ella es implicada por las claves candidatas.

REFERENCIAS

https://es.wikipedia.org/wiki/Normalizaci3n_de_bases_de_datos