

# Normalizacion

Rueda Rojas Luz Elena, Rojas Ruiz Luis Enrique

March 2020

## 1 ¿Qué es la normalización de bases de datos?

Es el proceso de organizar los datos de una base de datos. Debemos tener en cuenta la creación de tablas y las reglas que se usan para definir las relaciones, estas reglas son diseñadas para proteger los datos, y para que la base de datos sea flexible con el fin de eliminar redundancias y dependencias incoherentes. La normalización es la transformación de las vistas de usuario complejas y del almacén de datos a un juego de estructuras de datos más pequeñas y estables. Además de ser más simples y estables, las estructuras de datos son más fáciles de mantener que otras estructuras de datos.

## 2 ¿Por qué se normalizan las bases de datos?

‘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 los datos.
- Facilitar el acceso e interpretación de los datos.
- Reducir el tiempo y complejidad de revisión de las bases de datos.
- Optimizar el espacio de almacenamiento.
- Prevenir borrados indeseados de datos.

## 3 Requisitos de la normalización

Para que las tablas de nuestra BD estén normalizadas deben cumplir las siguientes reglas:

- 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.

## 4 Reglas o niveles de normalización

Para normalizar una base de datos existen principalmente 3 reglas, las cuales se deberían cumplir para evitar redundancias e incoherencias en las dependencias. A estas reglas se les conoce como "Forma normal" qué va de la 1 a la 3 y si la base de datos cumple con cada regla se dice que está en la "primera o segunda o tercera forma normal"

Aunque son posibles otros niveles de normalización, la tercera forma normal se considera el máximo nivel necesario para la mayoría de las aplicaciones.

## 5 Primera forma normal

"El valor de una columna debe ser una entidad atómica, indivisible, excluyendo así las dificultades que podría conllevar el tratamiento de un dato formado de varias partes"

Lo anterior lo podemos explicar en tres pasos:

- Eliminar los grupos repetidos de las tablas individuales.
- Crear una tabla independiente para cada conjunto de datos relacionados.
- Identificar cada conjunto de datos relacionados con una clave principal.
- Todos los atributos, valores almacenados en las columnas, deben ser indivisibles y no deben existir grupos de valores repetidos.

## 6 Segunda forma normal

Las tablas que están ajustadas a la primera forma normal, y además disponen de una clave primaria formada por una única columna con un valor indivisible, cumplen ya con la segunda forma normal. Ésta afecta exclusivamente a las tablas en las que la clave primaria está formada por los valores de dos o más columnas, debiendo asegurarse, en este caso, que todas las demás columnas son accesibles a través de la clave completa y nunca mediante una parte de esa clave. Para la segunda regla se debe:

- Crear tablas independientes para conjuntos de valores que se apliquen a varios registros.
- Relacionar estas tablas con una clave externa.

## 7 Tercera forma normal

En cuanto a la tercera forma normal, ésta indica que no deben existir dependencias transitivas entre las columnas de una tabla, lo cual significa que las columnas que no forman parte de la clave primaria deben depender sólo de la clave, nunca de otra columna no clave.

En esta regla se eliminan los campos que no dependen de la clave.

La normalización es una técnica utilizada para diseñar tablas en las que las redundancias de datos se reducen al mínimo. Las primeras tres formas normales (1FN, 2FN y 3FN) son las más utilizadas. Desde un punto de vista estructural, las formas de mayor nivel son mejores que las de menor nivel, porque aquellas producen relativamente pocas redundancias de datos en la base de datos. En otras palabras, 3FN es mejor que 2FN y ésta, a su vez, es mejor que 1FN. Casi todos los diseños de negocios utilizan la 3FN como forma ideal. en algunos casos, es necesario considerar la desnormalización para mejorar el rendimiento. La desnormalización es la duplicación intencionada de columnas en varias tablas, lo cual aumenta la redundancia de datos. No debe suponerse automáticamente que todas las uniones tardan demasiado tiempo.

## 8 Desnormalizacion

Las reglas de normalización no consideran el rendimiento. En algunos casos es necesario considerar la desnormalización para mejorar el rendimiento. La desnormalización es la duplicación intencionada de columnas en varias tablas, lo cual aumenta la redundancia de datos.

La normalización crea más tablas al avanzar hacia formas normales más altas, sin embargo, a mayor número de tablas, mayor número de combinaciones al recuperar los datos; lo que contribuye a la ralentización de las consultas. Por esta razón, para mejorar la velocidad de determinadas consultas, se pueden anular las ventajas de la integridad de datos y devolver la estructura de los datos a una forma normal inferior.

la unión de muchas tablas requiere operaciones de entrada/salida (I/O) y lógica de procesamiento adicional en el disco, con lo que se reduce la velocidad del sistema. Por lo tanto, pueden existir circunstancias fortuitas que permitan algún grado de desnormalización para incrementar la velocidad de procesamiento. Debemos tomar en cuenta que la ventaja de una mayor velocidad de procesamiento debe evaluarse cuidadosamente contra la desventaja de datos anómalos. a segunda forma normal añade la necesidad de que no existan dependencias funcionales parciales. Esto significa que todos los valores de las columnas de una fila deben depender de la clave primaria de dicha fila, entendiendo por clave primaria los valores de todas las columnas que la formen, en caso de ser más de una.

## References

Coronel, C., Morris, S. and Rob, P., 2011. Bases De Datos. México, D.F.: CENGAGE Learning.