

NORMALIZACIÓN

Luis Mauricio Barrientos Veana
Bases de datos

18 de marzo de 2020

1. Qué es 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. (Kendall, 2005).

La normalización de una base de datos persigue varios objetivos, principalmente reducir la redundancia de datos y simplificar las dependencias entre columnas (atributos), aplicándose de manera acumulativa. Esto quiere decir que la segunda forma normal incluye a la primera, la tercera a la segunda y así sucesivamente. Por lo tanto una base de datos que esté en segunda forma normal cumplirá las dos primeras reglas de normalización.

2. Primera forma normal (1FN)

La primera regla de normalización se expresa generalmente en forma de dos indicaciones separadas.

Todos los atributos, valores almacenados en las columnas, deben ser indivisibles. No deben existir grupos de valores repetidos.

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.

Supongamos que tienes en una tabla una columna Dirección para almacenar la dirección completa, dato que se compondría del nombre de la calle, el número exterior, el número interior (puerta), el código postal, el estado y la capital.

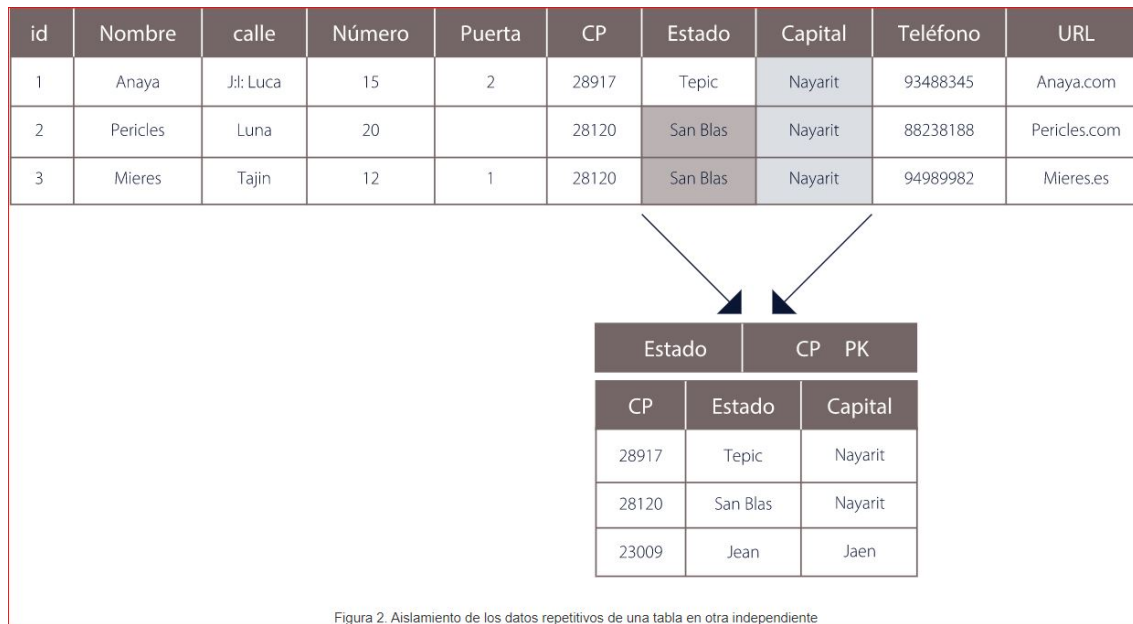


Una tabla con esta estructura plantea problemas a la hora de recuperar información. Imagina que necesitas conocer todas las entradas correspondientes a una determinada población, o que quieres buscar a partir del código postal. Al ser la dirección completa una secuencia de caracteres de estructura libre no resultaría nada fácil.

Existirán más columnas, pero cada una de ellas contendrá un valor simple e indivisible que facilitará la realización de las operaciones antes mencionadas.

En cuanto a la segunda indicación, se debe evitar la repetición de los datos de la población y provincia en cada una de las filas. Siempre que al muestrear la información de una tabla aparezcan datos repetidos, existe la posibilidad de crear una tabla independiente con ellos.

Si el diseño de nuestra base de datos cumple estas premisas, está preparada para pasar de la primera a la segunda forma normal.



3. Segunda forma normal (2FN)

Además de cumplir con las dos reglas del punto previo, la 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.

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.

4. Tercera forma normal (3FN)

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.

5. Referencias

M. Castañeda, "Normalización de Bases de Datos", Programas.cuaed.unam.mx, 2020. [On-line]. Available: https://programas.cuaed.unam.mx/repositorio/moodle/pluginfile.php/872/mod_resource/content/1/contenido/index.html. [Accessed : 15 – Mar – 2020].