

Capítulo 4.- Teoría de la Normalización

- 1.- INTRODUCCIÓN
- 2.- PRIMERA FORMA NORMAL
- 3.- SEGUNDA FORMA NORMAL
- 4.- TERCERA FORMA NORMAL

1.- INTRODUCCIÓN

El propósito de una Base de Datos es el de almacenar y facilitar información relativa a sujetos comunes que se agrupan formando estructuras.

El objeto de la normalización está basada en el concepto de las formas normales. Se dice que una estructura está en una forma normal determinada si satisface un cierto conjunto de restricciones.

2.- PRIMERA FORMA NORMAL (FN1)

Una estructura estará en FN1 si no contiene grupos repetitivos.

Supongamos que queremos conocer información acerca de los productos que realizan una serie de clientes en cada uno de sus pedidos y los datos que disponemos son los siguientes:

COD_CLI: Código del cliente que realiza el pedido

NOM_CLI: Nombre del cliente que realiza el pedido

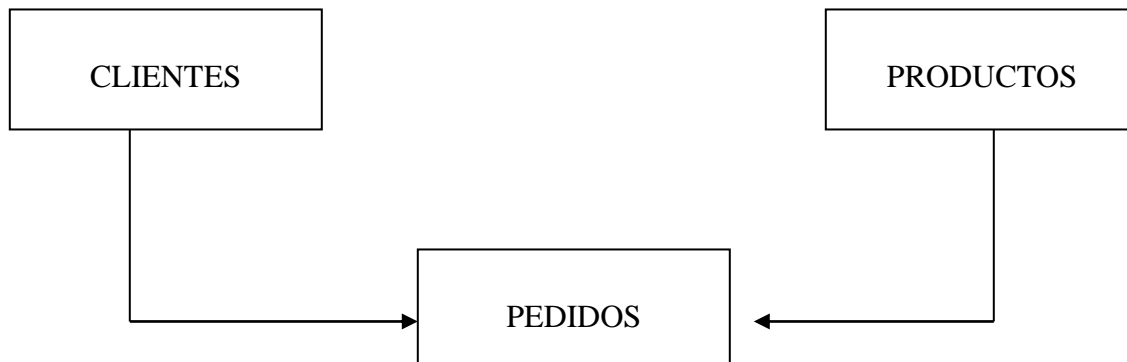
COD_POSTAL: Código postal del municipio donde vive el cliente

MUNICIPIO: Municipio donde vive el cliente

COD_PROD: Código del producto que pide el cliente

CANTIDAD: N° de unidades de producto que pide el cliente

PRECIO: Precio unitario del producto



La siguiente estructura está en FN1 siendo los campos clave COD_CLI y COD_PROD.

COD_CLI	NOM_CLI	COD_POSTAL	MUNICIPIO	COD_PROD	CANTIDAD	PRECIO
1	JUAN	45500	YEPES	P1	10	75
1	JUAN	45500	YEPES	P2	12	100
1	JUAN	45500	YEPES	P3	8	125
1	JUAN	45500	YEPES	P4	25	50
1	JUAN	45500	YEPES	P5	7	150
2	ANGEL	28300	FUENLA	P1	12	75
2	ANGEL	28300	FUENLA	P2	25	100
2	ANGEL	28300	FUENLA	P3	8	125
3	MARIA	08240	HOSPITA	P2	12	100
3	MARIA	08240	HOSPITA	P3	25	125
4	LUISA	14320	ALCAZAR	P3	25	125
4	LUISA	14320	ALCAZAR	P4	8	50
4	LUISA	14320	ALCAZAR	P5	12	150
5	JOSE	21380	MEDINA	P4	7	50

Esta estructura plantea problemas en actualizaciones:

a.- INSERCIÓN: No se puede reflejar que un cliente está ubicado en una ciudad hasta que no pida al menos un producto.

Por ejemplo, no se puede reflejar, solamente que el cliente 6 es de Madrid, pues obligamos a dar valores nulos a COD_PROD que es clave principal

b.- BORRADO: Si borramos una línea que sea la única de un cliente perderemos la información sobre el pedido y la ciudad donde vive dicho cliente.

c.- MODIFICACIÓN: Si queremos actualizar un dato repetido nos obliga a realizar búsquedas.

Por ejemplo, si el cliente 1 cambia su domicilio de YEPES a LILLO, nos obligará a modificar en las líneas que correspondan a dicho cliente YEPES por LILLO.

3.- SEGUNDA FORMA NORMAL (FN2)

Si la estructura anterior se divide en dos:

COD_CLI NOM_CLI COD_POSTAL MUNICIPIO

COD_CLI COD_PROD CANTIDAD PRECIO

La primera contiene información sobre el cliente, la segunda sobre los pedidos.

CLIENTE(campo clave COD_CLI)

COD_CLI NOM_CLI COD_POSTAL MUNICIPIO

1	JUAN	45500	YEPES
2	ANGEL	28300	FUENLA
3	MARIA	08240	HOSPITA
4	LUISA	14320	ALCAZAR
5	JOSE	21380	MEDINA

PEDIDO(campo clave COD_CLI + COD_PROD)

COD_CLI COD_PROD CANTIDAD PRECIO

1	P1	10	75
1	P2	12	100
1	P3	8	125
1	P4	25	50
1	P5	7	150
2	P1	12	75
2	P2	25	100
2	P3	8	125
3	P2	12	100
3	P3	25	125
4	P3	25	125
4	P4	8	50
4	P5	12	150
5	P4	7	50

Se resuelven los problemas anteriormente detectados en la actualización:

a.- INSERCIÓN: Podemos almacenar la información de un cliente aunque no pida ningún producto.

b.- BORRADO: Podemos borrar el pedido del cliente 5 sin perder la información de donde se encuentra dicho cliente

c.- MODIFICACIÓN: En la estructura del cliente la ciudad aparece una sola vez, por tanto bastará cambiar en la línea del cliente 1 YEPES por LILLO.

Con esto tenemos normalizada la estructura en FN2.

UNA ESTRUCTURA ESTÁ EN SEGUNDA FORMA NORMAL SI LO ESTÁ EN PRIMERA Y CADA CAMPO NO CLAVE DEPENDE DE TODA LA CLAVE.

La estructura CLIENTE está en FN2 porque los campos NOM_CLI COD_POSTAL y MUNICIPIO dependen de la clave COD_CLI. La estructura PEDIDO también lo está, puesto que los campos CANTIDAD PRECIO dependen de toda la clave **COD_CLI + COD_PROD**

4.- TERCERA FORMA NORMAL (FN3)

CLIENTE (campo clave COD_CLI)

	COD_CLI	NOM_CLI	COD_POSTAL	MUNICIPIO
--	---------	---------	------------	-----------

1		JUAN	45500	YEPES
2		ANGEL	28300	FUENLA
3		MARIA	08240	HOSPITA
4		LUISA	14320	ALCAZAR
5		JOSE	21380	MEDINA

Analizando la estructura CLIENTE encontramos que el campo COD_POSTAL depende del campo no clave MUNICIPIO, lo que origina los siguientes problemas:

a.- INSERCIÓN: No se puede asignar un Código Postal a un Municipio hasta que no tengamos un cliente en ella, debido a que no podemos dar un código a un municipio y no asignar valor a la clave COD_CLI

b.- BORRADO: Si borramos un cliente perderemos información del Código asignado al municipio en los casos en los que solo haya un cliente por municipio.

c.- MODIFICACIÓN: Si el código postal aparece repetido de forma que queremos cambiarlo, debemos buscar todas las líneas donde aparezca dicho código para cambiarlo.

La solución consiste en dividir esta estructura en dos:

COD_CLI NOM_CLI COD_POSTAL

COD_POSTAL MUNICIPIO

La primera estructura contiene información acerca de los CLIENTES y la segunda acerca de los MUNICIPIOS

CLIENTES (campo clave COD_CLI)

COD_CLI NOM_CLI COD_POSTAL

1	JUAN	45500
2	ANGEL	28300
3	MARIA	08240
4	LUISA	14320
5	JOSE	21380

MUNICIPIOS (campo clave COD_POSTAL)

COD_POSTAL MUNICIPIO

45500	YEPES
28300	FUENLA
08240	HOSPITA
14320	ALCAZAR
21380	MEDINA

De esta forma podremos:

a.- Asignar un código a un municipio sin tener un cliente de ese municipio.

b.- Borrar un cliente sin perder información del código asignado al municipio.

c.- Modificar el código del municipio en una sola línea.

CONCLUSIÓN:

UNA ESTRUCTURA ESTÁ EN TERCERA FORMA NORMAL SI LO ESTÁ EN SEGUNDA Y LOS CAMPOS QUE NO SON CLAVE SON INDEPENDIENTES ENTRE SÍ. O DICHO DE OTRA FORMA, LOS CAMPOS DEPENDEN DE TODA LA CLAVE Y SOLO DE LA CLAVE.

Por tanto nuestra estructura inicial

COD_CLI NOM_CLI COD_POSTAL MUNICIPIO COD_PROD CANTIDAD PRECIO

En tercera forma normal a quedado dividida en tres:

CLIENTES (COD_CLI NOM_CLI COD_POSTAL)

MUNICIPIOS (COD_POSTAL MUNICIPIO)

PEDIDOS (COD_CLI COD_PROD CANTIDAD PRECIO)