Normalización de Bases de Datos

Situación práctica

Formas normales

Serie de reglas que organizan la información de una base de datos de manera correcta. Deben ser tenidas en cuenta desde el diseño de la base de datos y dependen del contexto del problema a representar.

Primera forma normal	Todos los dominios subyacentes de los atributos contendrán valores atómicos.
Segunda forma normal	Todos los atributos no clave dependerán por completo de la clave primaria.
Tercera forma normal	Todos los atributos no clave dependen de manera no transitiva de la clave primaria.

1FN - Situación problemática

Contactos

DNI	Apellido	Nombres	Teléfonos
1000	Fernández	Carlos Martín	1111-1111, 2222-2222
2000	López	Juliana	3333-3333, 4444-4444, 5555-5555
3000	Freire	Dalmiro	No tiene
4000	Villalba	Julieta	6666-6666
5000	Perrone	Candela	7777-7777, 1111-1111

¿Cómo procedemos para eliminar el teléfono 4444-4444 del contacto López Juliana?

¿Cómo contabilizamos cuántos teléfonos tiene cada contacto?

1FN - ¿Solución verdadera?

Contactos

DNI	Apellido	Nombres	Teléfono 1	Teléfono 2	Teléfono 3
1000	Fernández	Carlos Martín	1111-1111	2222-2222	-
2000	López	Juliana	3333-3333	4444-4444	5555-5555
3000	Freire	Dalmiro	-	-	-
4000	Villalba	Julieta	6666-6666	-	-
5000	Perrone	Candela	7777-7777	1111-1111	

1FN - Creamos más tablas y las relacionamos entre sí

Contactos

DNI	Apellido	Nombres
1000	Fernández	Carlos Martín
2000	López	Juliana
3000	Freire	Dalmiro
4000	Villalba	Julieta
5000	Perrone	Candela

Teléfonos_x_Contacto

DNI	Teléfono
1000	1111-1111
1000	2222-2222
2000	3333-3333
2000	4444-4444
2000	5555-5555
4000	6666-6666
5000	7777-7777
1000	1111-1111

2FN - Situación problemática

Domicilios

DNI	Domicilio	Código Postal	Localidad
1000	Mitre 123	1111	Buenos Aires
2000	Belgrano 483	1111	Capital Federal
			Ciudad Autónoma
3000	Maipú 392	1111	de Buenos Aires
	Chacabuco		
4000	1020	4444	Tigre
5000	Arenales 8849	4444	Tigre

¿Qué ocurre con los nombres de Localidad de los tres primeros registros?

¿Cómo evitamos la repetición de datos en los nombres de Localidad?

¿De qué depende realmente el nombre de la Localidad?

2FN - Creamos tablas y relaciones para los campos que no dependen funcionalmente de la clave principal

Domicilios

DNI	Domicilio	Código Postal
1000	Mitre 123	1111
2000	Belgrano 483	1111
3000	Maipú 392	1111
4000	Chacabuco 1020	4444
5000	Arenales 8849	4444

Localidades

Código Postal	Localidad
1111	Buenos Aires
2222	General Pacheco
3333	San Fernando
4444	Tigre

3FN: Situación problemática

Contactos

DNI	Apellido	Nombres	IDArea	IDJefeArea
1000	Fernández	Carlos Martín	1	100
2000	López	Juliana	1	100
3000	Freire	Dalmiro	2	200
4000	Villalba	Julieta	2	200
5000	Perrone	Candela	3	300

Áreas

ID	Nombre	
1	Legales	
2	Sistemas	
3	Contable	

Jefes_x_Área

ID	Apellido	Nombre
100	Venezia	Pablo
200	Cerati	Eugenia
300	Monroe	Daniela

¿Cuál es el problema en este caso? Parece correcta. ¿No?

PISTA: Sólo hay un jefe por área.

3FN: Eliminar las columnas que dependen transitivamente de otro elemento

Contactos

DNI	Apellido	Nombres	IDArea
1000	Fernández	Carlos Martín	1
2000	López	Juliana	1
3000	Freire	Dalmiro	2
4000	Villalba	Julieta	2
5000	Perrone	Candela	3

Áreas

ID	Nombre	IDJefe
1	Legales	100
2	Sistemas	200
3	Contable	300

Jefes_x_Área

ID	Apellido	Nombre
100	Venezia	Pablo
200	Cerati	Eugenia
300	Monroe	Daniela