

Modelado de datos: Normalización

Ing. Luciano Straccia

TEMARIO



- ▶ Normalización
- ▶ Formas normales
 - ▶ 1FN
 - ▶ 2fn
 - ▶ 3fn

Anomalías - porqué normalizar

Normalización de bases de datos

Anomalías en las relaciones:

Anomalía de eliminación: Si al suprimir hechos de una entidad se suprimen hechos de otra entidad

ID_Estudiante	Actividad	Valor
1050	Futbol	2000
1100	Baloncesto	3000
1150	Natación	4000
1200	Futbol	2000

Si borramos al estudiante 1100 también se borra el hecho de que baloncesto cuesta 3000

Para registrar el hecho de que tejo cuesta 2000 habría que esperar a que alguien se inscriba

Anomalía de inserción: Si no podemos insertar un hecho sobre una entidad hasta que tengamos un hecho adicional de otra entidad

NORMALIZACION

Las formas normales son aplicadas a las tablas de una base de datos. Decir que una base de datos está en la forma normal **N** es decir que todas sus tablas están en la forma normal **N**.

En general, las primeras tres formas normales son suficientes para cubrir las necesidades de la mayoría de las bases de datos. El creador de estas 3 primeras formas normales (o reglas) fue Edgar Codd.

Veremos 1FN, 2FN y 3FN

NORMALIZACION



Primera forma
normal
1FN

NORMALIZACION - 1fn

Veamos este caso...

Código Cliente	Nombre	Apellido	Teléfono
123	Darío	Sanchez	4856-5545
456	Juan	Figueroa	02202-440020 4908-5656
789	Claudia	Martínez	4556-7878

Hay dos teléfonos para un mismo cliente. Y sabemos que no podemos tener 2 datos en un mismo campo...

NORMALIZACION - 1FN

¿qué hacemos? ¿esto?

Código Cliente	Nombre	Apellido	Teléfono	Teléfono
123	Darío	Sanchez	4856-5545	
456	Juan	Figueroa	02202-440020	4908-5656
789	Claudia	Martínez	4556-7878	

¿Y si pudiera haber 3 teléfonos? ¿Y si no supiéramos cuántos teléfonos?

NORMALIZACION - 1FN

Mejor esto... normalizamos.

Código Cliente	Nombre	Apellido
123	Darío	Sanchez
456	Juan	Figueroa
789	Claudia	Martínez

Código Cliente	Teléfono
123	4856-5545
456	02202-440020
456	4908-5656
789	4556-7878

Esto ahora está en 1era Forma Normal (1FN)

NORMALIZACION - 1FN

- Una relación de datos está en 1FN si cada dato contiene sólo un valor de su dominio o ninguno; si no existen repeticiones
- 1FN elimina grupos repetitivos

NORMALIZACION - 1FN

- Otra forma de verlo: en un Diccionario de Datos
- Si tengo

CLIENTE

Código

Nombre

Teléfono *(0..3)

- Claramente esto no está en 1FN y los teléfonos deben ir en tablas adicionales

SEGUNDA FORMA
NORMAL
2FN

NORMALIZACION - 2FN

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.

NORMALIZACION - 2FN

Veamos...

Clave

Alumno	Materia	Calificación	Nivel Curricular
Paola	Análisis Sistemas	7	2
Paola	Algebra	4	1
Adriana	DISEÑO Sistemas	10	3
Adriana	Algebra	10	1
Dario	Algebra	8	1

Hay un atributo no clave que depende sólo de una parte de la clave y no toda. No está en 2FN!

NORMALIZACION - 2FN

Solución: Normalizar

Alumno	Materia	Calificación
Paola	Análisis Sistemas	7
Paola	Algebra	4
Adriana	DISEÑO Sistemas	10
Adriana	Algebra	10
Dario	Algebra	8

Materia	Nivel Curricular
Análisis Sistemas	2
Algebra	1
DISEÑO Sistemas	3

NORMALIZACION - 1FN + 2FN

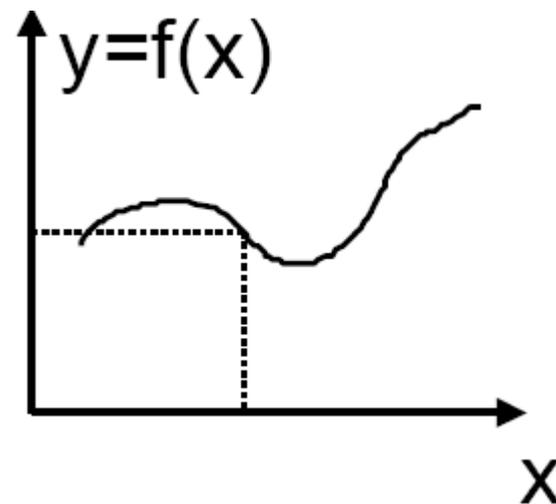
1FN: no hay grupos repetitivos

2FN: el atributo no clave depende de toda la clave primaria en vez de solo de una parte de ella.

Tercera forma
normal
3FN

DEPENDENCIA FUNCIONAL ENTRE ATRIBUTOS

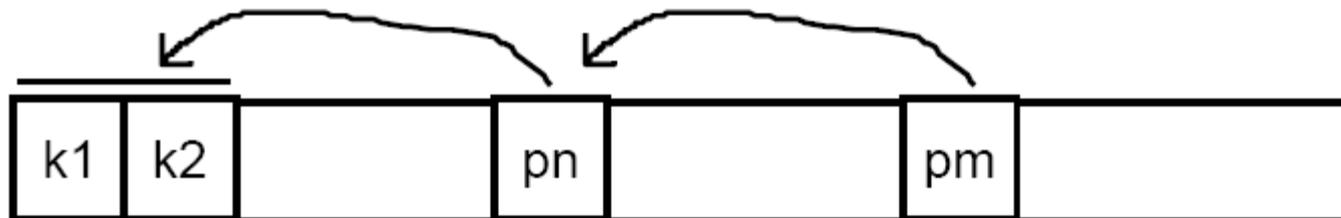
Sean X e Y subconjuntos de atributos de una relación. Diremos que Y tiene una dependencia funcional de X , o que X determina a Y , si cada valor de X tiene asociado siempre un único valor de Y .



DEPENDENCIA FUNCIONAL ENTRE ATRIBUTOS

Por ejemplo, si tenemos una relación con clientes de un hotel, y dos de sus atributos son el número de cliente y su nombre, podemos afirmar que el nombre tiene una dependencia funcional del número de cliente. Es decir, cada vez que en una tupla aparezca determinado valor de número de cliente, es seguro que el nombre de cliente será siempre el mismo..

DEPENDENCIA TRANSITIVA ENTRE ATRIBUTOS



DEPENDENCIA TRANSITIVA ENTRE ATRIBUTOS

Supongamos una tabla Ciudad con los atributos ciudad, población, superficie, renta, país, continente.

Población, superficie o renta tienen dependencia funcional de *ciudad*, así que de momento no nos preocupan.

DEPENDENCIA TRANSITIVA ENTRE ATRIBUTOS

En esta relación podemos encontrar también las siguientes dependencias:

Cada *ciudad* pertenece a un *país* y cada *país* a un *continente*, pero en cada *país* puede haber muchas *ciudades*. En este caso *continente* tiene una dependencia funcional transitiva con respecto a *ciudad*, a través de *país*.

NORMALIZACION - 3FN

Una tabla está en 3FN si está en 2FN y cada uno de sus atributos dependen funcional y directamente de la llave de la relación (o sea, no dependen transitoriamente de la misma).

Llave

↓

NombreCli depende de *IdPedido*
indirectamente a través de *IdCliente*

<u>IdPedido</u>	Fecha	IdCliente	NombreCli
123	23/11/1998	010	E.Metálicas
246	13/10/1998	020	M.Soldadura
280	5/12/1998	010	E.Metálicas

Pedido

NORMALIZACION - 1FN + 2FN + 3FN

1FN: no hay grupos repetitivos

2FN: 1FN + el atributo no clave depende de toda la clave primaria en vez de solo de una parte de ella.

3FN: 2FN + una tabla está en 3FN si está en 2FN y cada uno de sus atributos dependen funcional y directamente de la llave de la relación (o sea, no dependen transitoriamente de la misma).

FIN