

UNIVERSIDAD DE COSTA RICA

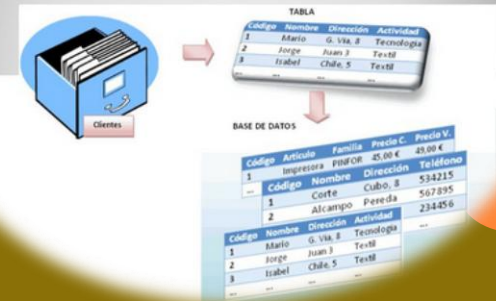


Normalización de las BASES DE DATOS

Curso: IF-4100 Fundamentos de las Bases de Datos

Profesor: MCI. Douglas Sanchez Artola

Normalización una BASE DE DATOS



Proceso de Normalización

Entendemos la normalización como el proceso que permite utilizar y diseñar bases de datos de manera óptima, evitando ambigüedades y errores de diseño, logrando un gran aprovechamiento de los recursos y un diseño preciso de la base de datos que se desee proyectar.



Para Korth & Sudarshan el objetivo principal de la normalización es:

“Generar un conjunto de esquemas de relaciones que permita almacenar información sin redundancia innecesaria, pero que también permita recuperar la información con facilidad.” (Silberschatz, Korth, & S.Sudarshan, 2006).



Problemas de un mal diseño de la base de datos

- Redundancia de datos
- Inconsistencia en los datos
- Ambigüedades en los datos
- Pérdida de la integridad de los datos
- Anomalías de operaciones de modificación de datos
- Uso inadecuado del almacenamiento
- Latencia en las consultas y transacciones



Objetivos de la normalización

- Eliminar redundancia de datos.
- Eliminar inconsistencia de los datos.
- Ordenar los datos.
- Eliminar errores lógicos.
- Optimizar el uso del espacio.
- Optimizar los tiempos de respuestas de las consultas y transacciones.
- etc



Tipos de Normalización

- Primera forma normal (1FN).
- Segunda forma normal (2FN).
- Tercera forma normal (3FN).
- Tercera forma normal Boyce Codd.
- Cuarta forma normal (4FN).
- Quinta forma normal (5FN)



Primera Forma Normal

- Establecidas por Edgar Frank Codd en los los 70`s
- La primera forma normal se encarga estrictamente de revisar que exista una relación entre los atributos y elementos que conforman la tabla, sin embargo también se encarga rigurosamente de que no existan valores repetidos.



Reglas que se deben cumplir

- Todos los atributos son atómicos. Un atributo es atómico si los elementos del dominio son indivisibles, mínimos.
- La tabla contiene una clave primaria única.
- La clave primaria no contiene atributos nulos.
- No debe existir variación en el número de columnas.
- Los campos no clave deben identificarse por la clave (Dependencia Funcional).
- Debe existir una independencia del orden tanto de las filas como de las columnas, es decir, si los datos cambian de orden no deben cambiar sus significados.
- Una tabla no puede tener múltiples valores en cada columna.
- Los datos son atómicos (a cada valor de X le pertenece un valor de Y y viceversa).



Ejemplo 1:

id	curso	slug	autor	materiales	fecha_creacion
1	Curso de Python	curso-de-python	David Aroesti	23	01/12/2017
2	Fundamentos de Bases de Datos	fundamentos-de-bases-de-datos	Pedro Díaz	15	02/12/2017
3	Curso de PostgreSQL	curso-de-postgresql	Mario Chávez	29	03/12/2017

Aplicando la Primera Forma Normal...

id	curso	slug	materiales	primer_nombre_profesor	segundo_nombre_profesor	fecha_creacion
1	Curso de Python	curso-de-python	23	David	Aroesti	01/12/2017
2	Fundamentos de Bases de Datos	fundamentos-de-bases-de-datos	15	Pedro	Díaz	02/12/2017
3	Curso de PostgreSQL	curso-de-postgresql	29	Mario	Chávez	03/12/2017

Segunda Forma Normal

- Obligatoriamente debe haberse aplicado la Primera Forma Normal(1FN)
- Es necesario establecer las dependencias funcionales

Dependencia Funcional.

- La segunda forma normal nos invita a eliminar la dependencia parcial que exista entre los atributos.
- Una dependencia parcial se da cuando un subgrupo de atributos -no llave, no depende totalmente de la llave primaria de la tabla.
- En la segunda forma normal tenemos que evaluar que todos los atributos no llave, dependen totalmente de la llave primaria, y de no parte de ella.
- La dependencia debe ser total, de la llave primaria



Aplicando Segunda Forma Normal...

	A	B	C	D	E	F	G	H
1	Segunda forma							
2								
3	<u>carnet</u>	<u>sigla</u>	nombre_1	nombre_2	apellido_1	apellido_2	descripcion_curso	
4	b45343	ma-1030	Geovanny		briceño	García	Estrategias de lectura ingles 1	
5	b54637	if-6201	maylon		perez	perez	informatica aplicada a los negocios	
6	b24576	fi-4100	sairo	jesus	aleman	Blanco	fundamentos de bases de datos	
7	PERSONA						CURSO	
8	<u>carnet</u>	nombre_1	nombre_2	apellido_1	apellido_2		<u>sigla</u>	descripcion_curso
9	b45343	Geovanny		briceño	García		ma-1030	Estrategias de lectura ingles 1
10	b54637	maylon	Josue	perez	Arguedas		if-6201	informatica aplicada a los negocios
11	b24576	sairo	jesus	aleman	Blanco		fi-4100	fundamentos de bases de datos
12								

Tercera Forma Normal

- Está en 2FN
- Ningún atributo no principal depende transitivamente de las claves (no existen dependencias entre atributos no principales)

Dependencia Funcional Transitiva

- Sean X, Y, Z tres atributos (o grupos de atributos) de la misma entidad.
- Si $X \rightarrow Y$ y $Y \rightarrow Z$ entonces $X \rightarrow Z$



Aplicando Tercera Forma Normal...

<u>Cédula</u>	Sigla_Curso	Nombre1	Nombre2	Ape1	Ape2	Nom_Curso
501800128	IF 3000	JUAN	PABLO	SOTO	AMEDER	DESARROLLO WEB
412526965	IF 4000	PEDRO	PABLO	COTO	SARSAN	ANALISIS Y DISEÑO
352632532	IF 5000	ROSA	MARIA	SOLIS	MORA	PROGRAMACIÓN MÓVIL
695623652	IF 7000	HEIDY	MARIA	MENDOZA	CASTRO	ESTRUCTURAS



Aplicando Tercera Forma Normal...

- Se dice que hay una DEPENDENCIA TRANSITIVA en el momento que detectemos que un atributo no llave, dependa de otro no llave.
- La dependencia transitiva siempre la voy a encontrar

Estudiantes

<u>Cédula</u>	Nombre1	Nombre2	Ape1	Ape2
501800128	JUAN	PABLO	SOTO	AMEDER
412526965	PEDRO	PABLO	COTO	SARSAN
352632532	ROSA	MARIA	SOLIS	MORA
695623652	HEIDY	MARIA	MENDOZA	CASTRO

Cursos

Sigla_Curso	Nom_Curso
IF 3000	DESARROLLO WEB
IF 4000	ANALISIS Y DISEÑO
IF 5000	PROGRAMACIÓN MÓVIL
IF 7000	ESTRUCTURAS

Aplicando Tercera Forma Normal...

Estudiantes

<u>Cédula</u>	Nombre1	Nombre2	Ape1	Ape2
501800128	JUAN	PABLO	SOTO	AMEDER
412526965	PEDRO	PABLO	COTO	SARSAN
352632532	ROSA	MARIA	SOLIS	MORA
695623652	HEIDY	MARIA	MENDOZA	CASTRO

Cursos

<u>Sigla_Curso</u>	Nom_Curso
IF 3000	DESARROLLO WEB
IF 4000	ANALISIS Y DISEÑO
IF 5000	PROGRAMACIÓN MÓVIL
IF 7000	ESTRUCTURAS

Cursos_Matriculados

<u>Id_Matricula</u>	Cédula	Sigla_Curso
1	501800128	IF 3000
2	412526965	IF 4000
3	352632532	IF 5000
4	695623652	IF 7000



FORMA NORMAL BOYCE-CODD

- La Forma Normal Boyce-Codd (Denominada por sus siglas en ingles como BCNF o FNBC) es una forma normal utilizada en la normalización de bases de datos. Es una adaptación vagamente más segura de lo establecido en la Tercera Forma Normal (3FN).
- Es una etapa en que se deben agrupar los datos por afinidad, formando tablas las cuales se relacionan entre si mediante campos comunes, una tabla se considera en esta forma si y sólo si cada determinante o atributo es una llave candidata.
- La forma normal de Boyce-Codd requiere que no existan dependencias funcionales no triviales de los atributos que no sean un conjunto de la clave candidata. En base de datos un atributo determinante es un atributo del que depende funcionalmente de manera completa algún otro atributo. Todo determinante es una clave candidata.

FORMA NORMAL BOYCE-CODD

- Un ejemplo típico para mostrar una tabla que, estando en 3FN, mantiene dependencias funcionales, puede ser una tabla que posee los atributos Código Postal, Dirección y Ciudad, deduciendo que a Ciudades diferentes le corresponden códigos postales distintos.

<u>CodigoPostal</u>	Direccion	Ciudad
1	C/ Las Flores N°17	Liberia
2	Av. Bolívar este N°72	Nicoya



FORMA NORMAL BOYCE-CODD

- En este caso hay dependencia entre el Código Postal y la Ciudad, ya que, conocido el Código Postal se puede conocer la Ciudad, y conocida la Dirección y la Ciudad, se conoce el Código Postal.

<u>CodigoPostal</u>	Direccion	Ciudad
1	C/ Las Flores Nº17	Liberia
2	Av. Bolívar este Nº72	Nicoya



FORMA NORMAL BOYCE-CODD

- Para transformar la tabla en una tabla en FNBC se crea una tabla de Códigos Postales y Ciudades, eliminando de la tabla original la Ciudad, obteniéndose dos tablas, una con los atributos Dirección y Código Postal y otra con el Código Postal y la Ciudad

<u>CodigoPostal</u>	Direccion
1	C/ Las Flores N°17
2	Av. Bolívar este N°72

<u>CodigoPostal</u>	Ciudad
1	Liberia
2	Nicoya



CUARTA FORMA NORMAL

- La cuarta forma normal (4NF) es una forma normal usada en la normalización de bases de datos.
- La 4NF se asegura de que las dependencias multivaluadas independientes estén correcta y eficientemente representadas en un diseño de base de datos.
- La 4NF es el siguiente nivel de normalización después de la forma normal de Boyce-Codd (BCNF).
- Una tabla con una dependencia multivaluada es una donde la existencia de dos o más relaciones independientes muchos a muchos causa redundancia; y es esta redundancia la que es suprimida por la cuarta forma normal.

CUARTA FORMA NORMAL

- Cada fila indica que un restaurante dado puede entregar una variedad dada de pizza a un área dada.
- Note que debido a que la tabla tiene una clave única y ningún atributo no-clave, no viola ninguna forma normal hasta el BCNF. Pero debido a que las variedades de pizza que un restaurante ofrece son independientes de las áreas a las cuales el restaurante envía, hay redundancia en la tabla: por ejemplo, nos dicen tres veces que A1 Pizza ofrece la Corteza rellena, y si A1 Pizza comienza a producir pizzas de Corteza de queso entonces necesitaremos agregar múltiples registros, uno para cada una de las Áreas de envío de A1 Pizza. En términos formales, esto se describe como que Variedad de pizza está teniendo una dependencia multivalor en Restaurante.

Permutaciones de envíos de pizzas

<u>Restaurante</u>	<u>Variedad de Pizza</u>	<u>Área de envío</u>
Vincenzo's Pizza	Corteza gruesa	Springfield
Vincenzo's Pizza	Corteza gruesa	Shelbyville
Vincenzo's Pizza	Corteza fina	Springfield
Vincenzo's Pizza	Corteza fina	Shelbyville
Elite Pizza	Corteza fina	Capital City
Elite Pizza	Corteza rellena	Capital City
A1 Pizza	Corteza gruesa	Springfield
A1 Pizza	Corteza gruesa	Shelbyville
A1 Pizza	Corteza gruesa	Capital City
A1 Pizza	Corteza rellena	Springfield
A1 Pizza	Corteza rellena	Shelbyville
A1 Pizza	Corteza rellena	Capital City

CUARTA FORMA NORMAL

- Para satisfacer la 4NF, debemos poner los registros sobre las variedades de pizza ofrecidas en una tabla diferente de los registros sobre áreas de envío:

Variedades por restaurante	
<u>Restaurante</u>	<u>Variedad de pizza</u>
Vincenzo's Pizza	Corteza gruesa
Vincenzo's Pizza	Corteza fina
Elite Pizza	Corteza fina
Elite Pizza	Corteza rellena
A1 Pizza	Corteza gruesa
A1 Pizza	Corteza rellena

Áreas de envío por restaurante	
<u>Restaurante</u>	<u>Área de envío</u>
Vincenzo's Pizza	Springfield
Vincenzo's Pizza	Shelbyville
Elite Pizza	Capital City
A1 Pizza	Springfield
A1 Pizza	Shelbyville
A1 Pizza	Capital City

CUARTA FORMA NORMAL

A partir de estas observaciones, se define la cuarta forma normal: Una afinidad está en cuarta forma normal si está en BCNF y no tiene dependencias de valores múltiples.



QUINTA FORMA NORMAL

Es un nivel de normalización de bases de datos diseñado para reducir redundancia en las bases de datos relacionales que guardan hechos multi-valores aislando semánticamente relaciones múltiples relacionadas. Una tabla se dice que está en 5NF si y sólo si está en 4NF y cada dependencia de unión (join) en ella es implicada por las claves candidatas.



QUINTA FORMA NORMAL

- La tabla está en 4FN
- No existen relaciones de dependencias de reunión (join) no triviales que no se generen desde las claves. Una tabla que se encuentra en la 4FN se dice que está en la 5FN si, y sólo si, cada relación de dependencia de reunión (join) se encuentra definida por claves candidatas. Por lo que si se aplicara una consulta entre al menos tres relaciones independientes entre sí dentro de la 4FN y se obtuvieran tuplas espurias, entonces no estaría dentro de la 5FN.



QUINTA FORMA NORMAL

Normalización

Dependencia de Unión Natural

Sea $R_1, R_2, R_3, \dots, R_n$ una descomposición de R

$$\text{Si } R_1 \bowtie R_2 \bowtie R_3 \dots \bowtie R_n = R$$

R satisface la Dependencia de Unión Natural

$$\alpha_n \rightarrow \beta_n \quad \alpha_n^+ \neq R \longrightarrow R \text{ no está en 5FN}$$

No trivial

QUINTA FORMA NORMAL. Ejemplo #1

Dependencia de Unión Natural

Empleado	Puesto	Proyecto
Andrés	Analista	Proyecto A
John	Programador	Proyecto B
Pedro	Instalador	Proyecto C
Ana	Diseñadora	Proyecto A

Empleado	Puesto
Andrés	Analista
John	Programador
Pedro	Instalador
Ana	Diseñadora

Empleado	Proyecto
Andrés	Proyecto A
John	Proyecto B
Pedro	Proyecto C
Ana	Proyecto A

Puesto	Proyecto
Analista	Proyecto A
Programador	Proyecto B
Instalador	Proyecto C
Diseñadora	Proyecto A

QUINTA FORMA NORMAL

Dependencia de Unión Natural

Empleado	Puesto
Andrés	Analista
John	Programador
Pedro	Instalador
Ana	Diseñadora

Empleado	Proyecto
Andrés	Proyecto A
John	Proyecto B
Pedro	Proyecto C
Ana	Proyecto A

Empleado	Puesto	Proyecto
Andrés	Analista	Proyecto A
John	Programador	Proyecto B
Pedro	Instalador	Proyecto C
Ana	Diseñadora	Proyecto A

Puesto	Proyecto
Analista	Proyecto A
Programador	Proyecto B
Instalador	Proyecto C
Diseñadora	Proyecto A

QUINTA FORMA NORMAL

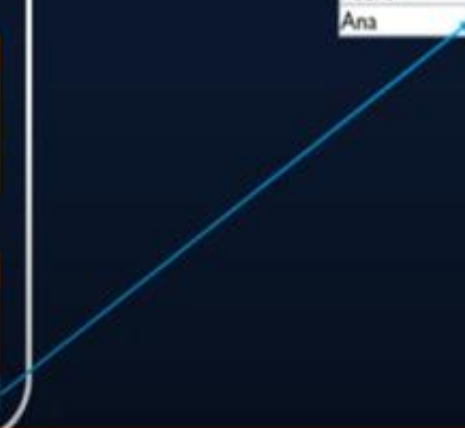
Dependencia de Unión Natural

Empleado	Puesto
Andrés	Analista
John	Programador
Pedro	Instalador
Ana	Diseñadora

Empleado	Proyecto
Andrés	Proyecto A
John	Proyecto B
Pedro	Proyecto C
Ana	Proyecto A

Puesto	Proyecto
Analista	Proyecto A
Programador	Proyecto B
Instalador	Proyecto C
Diseñadora	Proyecto A

Empleado	Puesto	Proyecto
Andrés	Analista	Proyecto A
John	Programador	Proyecto B
Pedro	Instalador	Proyecto C
Ana	Diseñadora	Proyecto A



QUINTA FORMA NORMAL

Dependencia de Unión Natural

Empleado	Puesto
Andrés	Analista
John	Programador
Pedro	Instalador
Ana	Diseñadora

Empleado	Proyecto
Andrés	Proyecto A
John	Proyecto B
Pedro	Proyecto C
Ana	Proyecto A

Puesto	Proyecto
Analista	Proyecto A
Programador	Proyecto B
Instalador	Proyecto C
Diseñadora	Proyecto A

Empleado	Puesto	Proyecto
Andrés	Analista	Proyecto A
John	Programador	Proyecto B
Pedro	Instalador	Proyecto C
Ana	Diseñadora	Proyecto A

=

Empleado	Puesto	Proyecto
Andrés	Analista	Proyecto A
John	Programador	Proyecto B
Pedro	Instalador	Proyecto C
Ana	Diseñadora	Proyecto A

$$R_1 \bowtie R_2 \bowtie R_3 = R$$

Hay una Dependencia de Unión Natural

R no está en 5NF

QUINTA FORMA NORMAL

No está en 5ta FN porque puede dividirse de una forma NO ADITIVA

Dependencia de Unión Natural

Empleado	Puesto	Proyecto
Andrés	Analista	Proyecto A
John	Programador	Proyecto B
Pedro	Instalador	Proyecto C
Ana	Diseñadora	Proyecto A

Empleado	Puesto
Andrés	Analista
John	Programador
Pedro	Instalador
Ana	Diseñadora

Empleado	Proyecto
Andrés	Proyecto A
John	Proyecto B
Pedro	Proyecto C
Ana	Proyecto A

Puesto	Proyecto
Analista	Proyecto A
Programador	Proyecto B
Instalador	Proyecto C
Diseñadora	Proyecto A

QUINTA FORMA NORMAL. Ejemplo #2

Dependencia de Unión Natural

Productora	Canal	Producto
Inter TV	Canal Norte	Series
Inter TV	Canal Norte	Novelas
Inter TV	Cine Canal	Películas
Film Stars	Canal Norte	Películas

Productora	Canal
Inter TV	Canal Norte
Inter TV	Canal Norte
Inter TV	Cine Canal
Film Stars	Canal Norte

Productora	Producto
Inter TV	Series
Inter TV	Novelas
Inter TV	Películas
Film Stars	Películas

Canal	Producto
Canal Norte	Series
Canal Norte	Novelas
Cine Canal	Películas
Canal Norte	Películas

QUINTA FORMA NORMAL

Dependencia de Unión Natural

Productora	Canal
Inter TV	Canal Norte
Inter TV	Cine Canal
Film Stars	Canal Norte

Productora	Producto
Inter TV	Seriales
Inter TV	Novelas
Inter TV	Peliculas
Film Stars	Peliculas

Canal	Producto
Canal Norte	Seriales
Canal Norte	Novelas
Canal Norte	Peliculas
Cine Canal	Peliculas

Productora	Canal	Producto
Inter TV	Canal Norte	Seriales
Inter TV	Canal Norte	Novelas
Inter TV	Canal Norte	Peliculas
Inter TV	Cine Canal	Seriales
Inter TV	Cine Canal	Novelas
Inter TV	Cine Canal	Peliculas
Film Stars	Canal Norte	Peliculas

QUINTA FORMA NORMAL

Dependencia de Unión Natural

Productora	Canal
Inter TV	Canal Norte
Inter TV	Cine Canal
Film Stars	Canal Norte

Productora	Producto
Inter TV	Series
Inter TV	Novelas
Inter TV	Películas
Film Stars	Películas

Canal	Producto
Canal Norte	Series
Canal Norte	Novelas
Canal Norte	Películas
Cine Canal	Películas

Productora	Canal	Producto	
Inter TV	Canal Norte	Series	✓
Inter TV	Canal Norte	Novelas	✓
Inter TV	Canal Norte	Películas	✓
Inter TV	Cine Canal	Series	✗
Inter TV	Cine Canal	Novelas	✗
Inter TV	Cine Canal	Películas	✓
Film Stars	Canal Norte	Películas	✓

QUINTA FORMA NORMAL

Este otro caso....

Dependencia de Unión Natural

Productora	Canal
Inter TV	Canal Norte
Inter TV	Cine Canal
Film Stars	Canal Norte

Productora	Producto
Inter TV	Seriales
Inter TV	Novelas
Inter TV	Películas
Film Stars	Películas

Canal	Producto
Canal Norte	Seriales
Canal Norte	Novelas
Canal Norte	Películas
Cine Canal	Películas

Productora	Canal	Producto
Inter TV	Canal Norte	Seriales
Inter TV	Canal Norte	Novelas
Inter TV	Canal Norte	Películas
Inter TV	Cine Canal	Películas
Film Stars	Canal Norte	Películas

QUINTA FORMA NORMAL

SI está en 5ta FN porque puede dividirse de una forma ADITIVA

Dependencia de Unión Natural

Productora	Canal
Inter TV	Canal Norte
Inter TV	Cine Canal
Film Stars	Canal Norte

Productora	Producto
Inter TV	Seriales
Inter TV	Novelas
Inter TV	Películas
Film Stars	Películas

Canal	Producto
Canal Norte	Seriales
Canal Norte	Novelas
Canal Norte	Películas
Cine Canal	Películas

Productora	Canal	Producto
Inter TV	Canal Norte	Seriales
Inter TV	Canal Norte	Novelas
Inter TV	Canal Norte	Películas
Inter TV	Cine Canal	Películas
Film Stars	Canal Norte	Películas

≠

Productora	Canal	Producto
Inter TV	Canal Norte	Seriales
Inter TV	Canal Norte	Novelas
Inter TV	Cine Canal	Películas
Cine Canal	Canal Norte	Películas

$$R_1 \bowtie R_2 \bowtie R_3 \neq R$$

Descomposición Aditiva

No hay una Dependencia de Unión Natural

R está en 5NF