

NORMALIZACIÓN DE LAS TABLAS (Primera forma normal)

Normalización y Primera Forma Normal

El proceso de normalización de bases de datos consiste en aplicar una serie de reglas a las relaciones obtenidas tras el paso del modelo entidad-relación al modelo relacional.

Las bases de datos relacionales se normalizan para:

- Evitar la redundancia de los datos.
- Evitar problemas de actualización de los datos en las tablas.
- Proteger la integridad de los datos.

La primera forma normal (1FN), requiere que los datos sean atómicos. En otras palabras, la 1FN **prohíbe a un campo contener más de un dato de su dominio de columna**. También exige que todas las tablas **deben tener una clave primaria**. Por último, indica que una tabla **no debe tener atributos que acepten valores nulos**.

Cuando no existe normalización, se presentan anomalías en la base de datos. Estos problemas que ocasionan problemas al momento de insertar, modificar o eliminar datos.

Ejemplos:

Ejemplos:

- Múltiples valores:

C_CLIENTE	N_CLIENTE	N_Apellido	#_TELEFONO
c0001	Juan	Perez	443-6523
c0002	Pedro	Moreno	274-5421, 235-6425
c0003	Jose	Quispe	542-5344

Multiples datos en número de teléfono

La forma correcta sería:

C_CLIENTE	N_CLIENTE	N_Apellido	#_TELEFONO
c0001	Juan	Perez	443-6523
c0002	Pedro	Moreno	274-5421
c0003	Jose	Quispe	542-5344

Número de teléfono normalizado

- Redundancia de datos:

<u>C_CLIENTE</u>	<u>N_CLIENTE</u>	<u>N_Apellido</u>	<u>#_TELEFONO</u>
c0001	Juan	Perez	443-6523
c0002	Pedro	Moreno	274-5421
c0002	Pedro	Moreno	235-6425
c0003	Jose	Quispe	542-5344

Datos redundantes en dos registros

La forma correcta de representar la tabla sería:

<u>C_CLIENTE</u>	<u>N_CLIENTE</u>	<u>N_Apellido</u>	<u>C_CLIENTE</u>	<u>#_TELEFONO</u>
c0001	Juan	Perez	c0001	443-6523
c0002	Pedro	Moreno	c0002	274-5421
c0003	Jose	Quispe	c0002	235-6425
			c0003	542-5344

Sin redundancia. Cabe mencionar que la llave primaria de la segunda tabla es compuesta

- Columnas que permiten valores nulos:

<u>C_CLIENTE</u>	<u>N_CLIENTE</u>	<u>N_Apellido</u>	<u>#_TELEFONO</u>	<u>#_TELEFONO2</u>
c0001	Juan	Perez	443-6523	
c0002	Pedro	Moreno	274-5421	235-6425
c0003	Jose	Quispe	542-5344	

La forma correcta de representar esta tabla sería como en el ejemplo anterior

<u>C_CLIENTE</u>	<u>N_CLIENTE</u>	<u>N_Apellido</u>	<u>C_CLIENTE</u>	<u>#_TELEFONO</u>
c0001	Juan	Perez	c0001	443-6523
c0002	Pedro	Moreno	c0002	274-5421
c0003	Jose	Quispe	c0002	235-6425
			c0003	542-5344

- Tabla sin llave principal:

<u>N_CLIENTE</u>	<u>N_Apellido</u>	<u>#_TELEFONO</u>
Juan	Perez	443-6523
Pedro	Moreno	274-5421
Jose	Quispe	542-5344

La forma correcta sería agregando una llave principal

<u>C_CLIENTE</u>	<u>N_CLIENTE</u>	<u>N_Apellido</u>	<u>#_TELEFONO</u>
c0001	Juan	Perez	443-6523
c0002	Pedro	Moreno	274-5421
c0003	Jose	Quispe	542-5344

SEGUNDA FORMA NORMAL

Segunda Forma Normal en Bases de Datos (2FN)

La segunda forma normal (2NF) es una forma normal usada en normalización de bases de datos. Una tabla que está en la primera forma normal (1NF) debe satisfacer criterios adicionales para calificar para la segunda forma normal.

Una tabla 1NF estará 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 una parte de ella.

Ejemplos:

- Ejemplo1:

<u>C TRABAJADOR</u>	<u>C AREA</u>	N TRABAJADOR	Q HORAS
c0001	a0023	Ernesto	45
c0002	a0034	Jhonatan	26
c0003	a0019	Antonio	68

Esto se soluciona separando el atributo N_TRABAJADOR a una tabla separada

<u>C TRABAJADOR</u>	<u>C AREA</u>	Q HORAS
c0001	a0023	45
c0002	a0034	26
c0003	a0019	68

C_TRABAJADOR	N_TRABAJADOR
c0001	Ernesto
c0002	Jhonatan
c0003	Antonio

- Ejemplo 2:

C_CURSO	C_ALUMNO	Q_NOTA	N_CURSO	N_ALUMNO
c0001	a0023	45	contabilidad	carlos
c0002	a0034	26	estadística	roberto
c0003	a0019	68	álgebra	frank

en este caso se puede separar la tabla en 3 tablas diferentes para cumplir con la 2FN

C_CURSO	N_CURSO	C_ALUMNO	N_ALUMNO
c0001	contabilidad	a0023	carlos
c0002	estadística	a0034	roberto
c0003	álgebra	a0019	frank

C_CURSO	C_ALUMNO	Q_NOTA
c0001	a0023	45
c0002	a0034	26
c0003	a0019	68

- Ejemplo3:

C_CIUADAD	C_COMPTRANSPORTE	Q_VEHICULOS	Q_CHOFRERES
c0001	t0345	23	30
c0002	t0741	27	32
c0003	t0184	25	33

Esta tabla esta en FN2

Ejemplo4:

C_PROFESOR	C_AULA	Q_PROMEDIO	Q_ALUMNOS
PCSUICOR	SI204	13	37
PCSUSIA	SW223	15	38
PCSIGALV	SI104	14	39

Esta tabla esta en FN2

La Segunda Forma Normal presenta anomalías, en donde si existe dependencia funcional completa entre los atributos. como en el ejemplo siguiente:

Torneo	Año	Ganador	Fecha de nacimiento del ganador
Des Moines Masters	1998	Chip Masterson	14 de marzo de 1977
Indiana Invitational	1998	Al Fredrickson	21 de julio de 1975
Cleveland Open	1999	Bob Albertson	28 de septiembre de 1968
Des Moines Masters	1999	Al Fredrickson	21 de julio de 1975
Indiana Invitational	1999	Chip Masterson	14 de marzo de 1977

De este tipo de anomalías se encarga la tercera forma normal.

TERCERA FORMA NORMAL

Tercera Forma Normal en Bases de Datos (3FN)

La tercera forma normal (3NF) es una forma normal usada en la normalización de bases de datos. Para que la tabla esté en 3NF, se debe cumplir que la tabla ya esté en 2NF. Adicionalmente, ningún atributo no primario de la tabla debe ser dependiente transitivamente de una clave primaria.

Ejemplos:

Torneo	Año	Ganador	Fecha de nacimiento del ganador
Des Moines Masters	1998	Chip Masterson	14 de marzo de 1977
Indiana Invitational	1998	Al Fredrickson	21 de julio de 1975
Cleveland Open	1999	Bob Albertson	28 de septiembre de 1968
Des Moines Masters	1999	Al Fredrickson	21 de julio de 1975
Indiana Invitational	1999	Chip Masterson	14 de marzo de 1977

esto se soluciona removiendo la transitividad y separando en dos relaciones diferentes

Torneo	Año	Ganador
Des Moines Masters	1998	Chip Masterson
Indiana Invitational	1998	Al Fredrickson
Cleveland Open	1999	Bob Albertson
Des Moines Masters	1999	Al Fredrickson
Indiana Invitational	1999	Chip Masterson

Jugador	Fecha de nacimiento del ganador
Chip Masterson	14 de marzo de 1977
Al Fredrickson	21 de julio de 1975
Bob Albertson	28 de septiembre de 1968
Al Fredrickson	21 de julio de 1975
Chip Masterson	14 de marzo de 1977

La anomalía que presenta la tercera forma normal no está en la inserción, modificación, o eliminación de datos. Sino, en la dependencia multivaluada.

Restaurante	Variedad de Pizza	Área de envío
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

Cuarta Forma Normal en Bases de Datos (4FN)

La cuarta forma normal (4NF) es una forma normal usada en la normalización de bases de datos. Esta se asegura que todas las dependencias multivaluadas esten representadas de la manera mas eficiente en la base de datos. Esta forma normal se realiza despues de la forma normal de Boyce-Codd.

Una dependencia multivaluada es una restricción completa entre dos grupos de atributos en una relación.

<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

Esta dependencia multivaluada se resuelve de la siguiente manera:

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

Hay algunas ocasiones en que aun existen anomalías después de una FN4, aunque este tipo de anomalías son del tipo semánticas. un ejemplo seria la siguiente tabla con su descripción:

<u>Psiquiatra</u>	<u>Asegurador</u>	<u>Condición</u>
Dr. James	Healthco	Ansiedad
Dr. James	Healthco	Depresión
Dr. Kendrick	FriendlyCare	OCD
Dr. Kendrick	FriendlyCare	Ansiedad
Dr. Kendrick	FriendlyCare	Depresión
Dr. Lowenstein	FriendlyCare	Esquizofrenia
Dr. Lowenstein	Healthco	Ansiedad
Dr. Lowenstein	Healthco	Demencia
Dr. Lowenstein	Victorian Life	Trastorno de conversión

Cuando un psiquiatra es autorizado a ofrecer el tratamiento reembolsable a los pacientes asegurados por el asegurador P, y el psiquiatra puede tratar la condición C, entonces - en caso que el asegurador P cubra la condición C - debe ser cierto que el psiquiatra puede ofrecer el tratamiento reembolsable a los pacientes que sufren de la condición C y están asegurados por el asegurador P.

Se puede dividir con esta descripción a la relación en tres partes:

Psiquiatra	Condición	Psiquiatra	Asegurador	Asegurador	Condición
Dr. James	Ansiedad	Dr. James	Healthco	Healthco	Ansiedad
Dr. James	Depresión	Dr. Kendrick	FriendlyCare	Healthco	Depresión
Dr. Kendrick	OCD	Dr. Lowenstein	FriendlyCare	Healthco	Demencia
Dr. Kendrick	Ansiedad	Dr. Lowenstein	Healthco	FriendlyCare	OCD
Dr. Kendrick	Depresión	Dr. Lowenstein	Victorian Life	FriendlyCare	Ansiedad
Dr. Lowenstein	Esquizofrenia			FriendlyCare	Depresión
Dr. Lowenstein	Ansiedad			FriendlyCare	Trastorno emocional
Dr. Lowenstein	Demencia			FriendlyCare	Esquizofrenia
Dr. Lowenstein	Trastorno de conversión			Victorian Life	Trastorno de conversión

Esta disposición ayuda a reducir redundancia y a facilitar cualquier tipo de búsquedas.

Información obtenida de:

<http://basesdedatosjc.blogspot.com/2012/04/primer-forma-normal-en-bases-de-datos.html>