

BASES DE DATOS: GENERALIDADES Y SISTEMAS DE GESTIÓN



NORMALIZACIÓN Y EL MODELO
ENTIDAD RELACIÓN (E-R)



NORMALIZACIÓN Y EL MODELO ENTIDAD RELACIÓN (E-R)

PRESENTACIÓN

La Normalización busca en su esencia evitar al máximo la redundancia de datos o inconsistencia de los mismos; por lo que conlleva a desarrollar el Diagrama Entidad-Relación, el cual está definido dentro de la Teoría de Bases de Datos Relacionales, para que sea una guía en la estructuración del diseño de las tuplas que están contenidas lógicamente en una entidad o tabla, y teniendo presente que el usuario final es quien va a trabajar la información. Es por eso que el diseño busca evitar que éste cometa errores que lleguen a afectar la información almacenada.

Por tal motivo, en este contenido de estudio se presenta qué es la Normalización, sus objetivos y ventajas, así como qué es el Modelo Entidad-relación, con el objetivo de tener los conceptos necesarios para asimilar la fundamentación y manejo de la construcción de Bases de Datos Relacionales.

Resultados de Aprendizaje: Describir las características de la Base de Datos requerida por una aplicación.



NORMALIZACIÓN Y EL MODELO ENTIDAD RELACIÓN (E-R)

TABLA DE CONTENIDO

PRESENTACIÓN	2
ÍNDICE DE FIGURAS	4
1. LA NORMALIZACIÓN	5
1.1. Reglas de Normalización.....	7
1.1.1. Primera forma normal (1NF).....	8
1.1.2. Segunda forma normal (2NF).....	12
1.1.3. Tercera forma normal (3NF).....	13
2. MODELO ENTIDAD RELACIÓN (E-R).....	16
GLOSARIO.....	18
BIBLIOGRAFÍA.....	19
WEBGRAFÍA.....	20
CRÉDITOS.....	21





NORMALIZACIÓN Y EL MODELO ENTIDAD RELACIÓN (E-R)

ÍNDICE DE FIGURAS

Figura 1. Objetivos de la normalización.....	6
Figura 2. Ventajas de la normalización.....	6
Figura 3. Reglas de la normalización	8



NORMALIZACIÓN Y EL MODELO ENTIDAD RELACIÓN (E-R)

1. LA NORMALIZACIÓN

La primera forma que se piensa para ordenar la información, es crear un solo sitio en dónde agruparla. Esto es válido si el manejo de información es pequeño, pero cuando va creciendo y se requiere que su manipulación genere resultados más eficientes, es necesario tener reglas o normas que permitan estructurar el almacenamiento de la información en Bases de datos. Estas reglas se conocen con el nombre de **normalización**, las cuales permiten que los diseñadores de Bases de datos, tengan una herramienta eficaz para su construcción.

Cabe aclarar, que la normalización es un concepto que permite generar el Modelo relacional de la BD, pero no es una técnica para el diseño de BD, es decir, un diseñador podría generar un diseño de Bases de datos sin normalización (equivale a construir una casa sin planos), pero no se puede aplicar un Modelo relacional sin normalización (equivale a elaborar planos y construir la casa de acuerdo a lo planteado en ellos). La normalización trabaja de forma ascendente, es decir, toma como base los atributos y los agrupa, de acuerdo a sus características en tablas relacionadas.

Por ende, con la aplicación de la normalización se determina la estructura de los datos sin importar la complejidad de la Base de datos, es decir, funciona igual para una Base de datos pequeña o grande, logrando simplificar su estructura. Además busca un muy buen diseño y procesamiento eficaz de los datos.

Por lo tanto los objetivos de la normalización son:



3

NORMALIZACIÓN Y EL MODELO ENTIDAD RELACIÓN (E-R)



Figura 1. Objetivos de la normalización
Fuente: SENA

Y sus ventajas son:

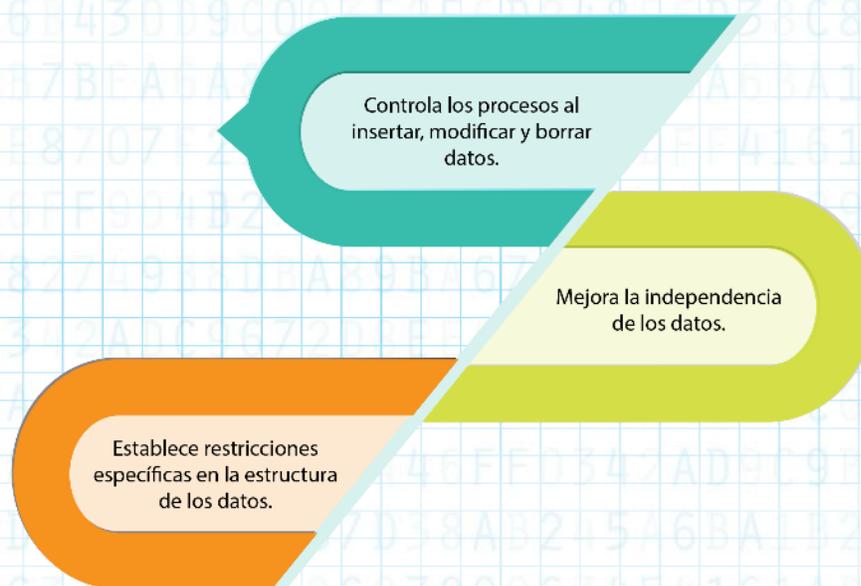


Figura 2. Ventajas de la normalización
Fuente: SENA



NORMALIZACIÓN Y EL MODELO ENTIDAD RELACIÓN (E-R)

Entonces, en resumen se debe normalizar para:

- Estructura los datos donde exista una relación entre ellos.
- Todas las acciones, consultas y reportes de los usuarios sean ágiles y de respuesta oportuna, haciendo sencillo el manejo de la BD.
- Facilita el mantenimiento de los datos.
- Posibilita la creación de nuevos atributos o datos cuando se generan nuevas necesidades o aplicaciones.

1.1. Reglas de Normalización

La normalización está dividida básicamente en tres niveles, Primera Forma Normal (1NF), Segunda Forma Normal (2NF) Y Tercera Forma Normal (3NF), cada una de ellas con sus propias reglas.

Existen BD que no requieren la aplicación de todas las formas de normalización, aunque las más complejas son las que llegan a 3NF.

Para una mayor comprensión la siguiente figura explica de forma simple la descripción de cada forma normal y más adelante se presentan ejemplos donde se aplica cada una de éstas.



NORMALIZACIÓN Y EL MODELO ENTIDAD RELACIÓN (E-R)



Figura 3. Reglas de la normalización
Fuente: SENA

1.1.1. Primera forma normal (1FN)

Su objetivo es:

- Dejar todos los campos con valores atómicos en las tablas individuales.
- Crear estos campos repetidos en una nueva tabla.
- Relacionar estos campos con la llave primaria de la primera tabla.



NORMALIZACIÓN Y EL MODELO ENTIDAD RELACIÓN (E-R)

Entonces, una entidad está en primera forma normal, si y solo si todos sus atributos son atómicos, es decir, aquellos que no pueden descomponerse en datos más simples. Es importante no crear campos para almacenar datos similares.

Ejemplo:

Documento	Cod_Ciudad	Ciudad	Nombres	Apellidos	Telefono_movil	Email	Sexo
1234	12	Ibagué	Diego	Silva	31234567	Disilva@msn.com diegosilva@hotmail.es	M
3466	24	Espinal	Andrea	Amaya	31345678; 31678990	Andreaa@jj.com	F
8896	18	Cali	Milena	Lozano	31845678	mlozano@hotmail.es	F
4345	12	Ibagué	Diego	Lozano	30023567; 31234578	dlozano@yahoo.es; diegolozano@gmail.com	M

Al aplicar la 1FN, se encuentra lo siguiente:

El atributo Telefono_movil, no posee valores atómicos, porque existen dos registros con 2 teléfonos en el mismo campo.

El atributo Email, no posee valores atómicos, porque existen dos registros con 2 correos en el mismo campo.

Entonces, atendiendo las reglas de la 1FN, generar la siguiente solución viola la 1FN.

Documento	Cod_Ciudad	Ciudad	Nombres	Apellidos	Telefono_movil	Telefono_movil 2	Email	Email 2	Sexo	Profesion
1234	12	Ibagué	Diego	Silva	31234567		Disilva@msn.com	diegosilva@hotmail.es	M	Ingeniero
3466	24	Espinal	Andrea	Amaya	31345678	31678990	Andreaa@jj.com		F	Arquitecto
8896	18	Cali	Milena	Lozano	31845678		mlozano@hotmail.es		F	Abogado
4345	12	Ibagué	Diego	Lozano	30023567	31234578	dlozano@yahoo.es;	diegolozano@gmail.com	M	Obrero





NORMALIZACIÓN Y EL MODELO ENTIDAD RELACIÓN (E-R)

Como se observa, se crearon dos campos de Telefono_movil_ y dos de Email, pero precisamente esto es lo que la 1FN indica que no se debe hacer. La razón principal es porque se va a perder mucho espacio en valores nulos, tal como se nota en el registro de Milena Lozano, ya que no utiliza los dos nuevos campos.

Entonces, aplicando la 1FN (todos los campos tienen valores atómicos), la tabla queda de la siguiente forma:

Documento	Cod_Ciudad	Ciudad	Nombres	Apellidos	Telefono_movil_	Email	Sexo	Profesion
1234	12	Ibagué	Diego	Silva	31234567	Disilva@msn.com	M	Ingeniero
1234	12	Ibagué	Diego	Silva	31234567	diegosilva@hotmail.es	M	Ingeniero
3466	24	Espinal	Andrea	Amaya	31345678	Andreaa@jj.com	F	Arquitecto
3466	24	Espinal	Andrea	Amaya	31345678	Andreaa@jj.com	F	Arquitecto
8896	18	Cali	Milena	Lozano	31845678	mlozano@hotmail.es	F	Abogado
4345	12	Ibagué	Diego	Lozano	30023567	diegolozano@yahoo.es	M	Obrero
4345	12	Ibagué	Diego	Lozano	30023567	dlozano@yahoo.es	M	Obrero
4345	12	Ibagué	Diego	Lozano	30023567	dlozano@yahoo.es	M	Obrero
4345	12	Ibagué	Diego	Lozano	30023567	dlozano@yahoo.es	M	Obrero

Por consiguiente, se determina que todos los atributos están en valores atómicos, aunque existen valores repetidos porque, por ejemplo, Diego Silva, Andrea Amaya y Diego Lozano tienen diferentes valores para los atributos Telefono_movil_ y Email.

Entonces, aplicando la 1FN (eliminar datos redundantes), las tablas quedan de la siguiente forma:



3

NORMALIZACIÓN Y EL MODELO ENTIDAD RELACIÓN (E-R)

PERSONAS

Documento (PK)	Cod_Ciudad	Ciudad	Nombres	Apellidos	Sexo	Profesion
1234	12	Ibagué	Diego	Silva	M	Ingeniero
3466	24	Espinal	Andrea	Amaya	F	Arquitecto
8896	18	Cali	Milena	Lozano	F	Abogado
4345	12	Ibagué	Diego	Lozano	M	Obrero

Luego se crean dos nuevas tablas con los campos repetidos y relacionados con la llave principal que es **Documento**.

MOVIL

Documento (PK)	Telefono_movil
1234	31234567
3466	31345678
3466	31678990
8896	31845678
4345	30023567
4345	31234578

CORREO

Documento (PK)	Email
1234	Disilva@msn.com
3466	diegosilva@hotmail.es
3466	Andreaa@jj.com
8896	mlozano@hotmail.es
4345	dlozano@yahoo.es
4345	diegolozano@gml.com

Por lo tanto, se determina que todos los campos son atómicos y que no existen valores repetidos, pero todos están relacionados con la llave principal (PK).



NORMALIZACIÓN Y EL MODELO ENTIDAD RELACIÓN (E-R)

1.1.2. Segunda forma normal (2NF)

Los objetivos principales de la segunda forma normal son:

- Determinar cuáles columnas que no son llave, no dependen de la llave primaria de la tabla.
- Eliminar dichas columnas de la tabla base.
- Crear una segunda tabla con dichas columnas, y la(s) columna(s) de la PK de la cual dependen (García, 2009).

Entonces en las tablas se tiene lo siguiente:

- Las tablas Movil y Correo, cumplen con la 2NF, debido a que sus atributos dependen directamente de la llave principal.
- La tabla Personas, posee dos campos que no son dependientes de la PK Documento, los cuales son: Cod_Ciudad y Ciudad. Aunque a cada persona se le expide su documento en una ciudad, no existe relación directa entre la ciudad y la persona; por consiguiente se debe separar en una nueva tabla, quedando de la siguiente forma:





NORMALIZACIÓN Y EL MODELO ENTIDAD RELACIÓN (E-R)

PERSONAS

Documento (PK)	Cod_Ciudad	Nombres	Apellidos	Sexo	Profesion
1234	12	Diego	Silva	M	Ingeniero
3466	24	Andrea	Amaya	F	Arquitecto
8896	18	Milena	Lozano	F	Abogado
4345	12	Diego	Lozano	M	Obrero

CIUDADES

Cod_Ciudad (PK)	Ciudad
12	Ibagué
24	Espinal
18	Cali

MOVIL

Documento (PK)	Telefono_movil
1234	31234567
3466	31345678
3466	31678990
8896	31845678
4345	30023567
4345	31234578

CORREO

Documento (PK)	Email
1234	Disilva@msn.com
1234	diegosilva@hotmail.com
3466	Andreaa@jj.com
8896	mlozano@hotmail.com
4345	dlozano@yahoo.es
4345	diegolozano@gml.com

Luego de aplicar la 2FN, se encuentra que la tabla **Ciudades** se utiliza en la tabla **Personas**, pero no es dependiente de ella. Por consiguiente, si se crean más personas, éstas pueden tener ciudades que ya existan. En caso de que no existan, se crean en la tabla **Ciudades**, pero no es necesario tener registrada la misma ciudad muchas veces.

1.1.3. Tercera forma normal (3NF)

Los objetivos principales de la tercera forma normal son:

- Determinar las columnas que son dependientes de otra columna no llave.
- Eliminar esas columnas de la tabla base.
- Crear una segunda tabla con esas columnas y con la columna no llave, de la cual son dependientes.



3

NORMALIZACIÓN Y EL MODELO ENTIDAD RELACIÓN (E-R)

Por ejemplo:

PERSONAS

Documento (PK)	Cod_Ciudad	Nombres	Apellidos	Sexo	Cod_Profesion
1234	12	Diego	Silva	M	10
3466	24	Andrea	Amaya	F	01
8896	18	Milena	Lozano	F	08
4345	12	Diego	Lozano	M	20

CIUDADES

Cod_Ciudad (PK)	Ciudad
12	Ibagué
24	Espinal
18	Cali

PROFESIONES

Cod_Profesion (PK)	Profesion
10	Ingeniero
01	Arquitecto
08	Abogado
20	Obrero



NORMALIZACIÓN Y EL MODELO ENTIDAD RELACIÓN (E-R)

MOVIL

Documento	Telefono_movil
(PK)	
1234	31234567
3466	31345678
3466	31678990
8896	31845678
4345	30023567
4345	31234578

CORREO

Documento	Email
(PK)	
1234	Disilva@msn.com
1234	diegosilva@hotmail.es
3466	Andreaa@jj.com
8896	mlozano@hotmail.es
4345	dlozano@yahoo.es
4345	diegolozano@gml.com

Cuando las tablas están en la 3NF se previenen errores de lógica, ya que cuando se inserta no se borran registros. Es decir, que cada columna en una tabla está identificada de manera única por la llave primaria, por lo que no deben existir datos repetidos. Esto proporciona un esquema limpio y elegante, fácil de trabajar y expandir.

Entonces, como se puede evidenciar, establecer y normalizar las Bases de datos, permiten asegurar la sincronización, almacenamiento, seguridad e integridad de la información, evitando con ello una dependencia estricta a una persona o empresa que contenga la estructura de la información. Además, generan un proceso ético-laboral al evitar esa exclusividad en el manejo de la arquitectura de la información, ya que le permite a la empresa realizar actualizaciones, adiciones y mantenimiento de sus Bases de datos, incluso pueden hacer migración a otro tipo de Base de datos sin perder su esencia, claridad y la responsabilidad de la información.

Por lo tanto, una vez se realice la normalización se organizará de manera visual la estructura de la Base de datos; a lo que se conoce con el nombre de **Diagrama Entidad-Relación**, tema que se tratará a continuación.



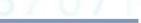
NORMALIZACIÓN Y EL MODELO ENTIDAD RELACIÓN (E-R)

2. MODELO ENTIDAD RELACIÓN (E-R)

El Modelo Entidad-Relación (E-R) es un modelo de datos que representa la abstracción y descripción de un sistema de información, el cual está conformado por entidades y sus relaciones, representándolo de forma gráfica y denominándose **diagrama entidad-relación**.

Para una empresa es fundamental diseñar el diagrama entidad-relación porque en él se encuentra de manera gráfica y formalizada toda la estructura de la Base de datos, para que luego no haya una dependencia específica hacia una persona o una empresa en el manejo de la Base de datos. Este aspecto es importante tenerlo presente en el ámbito laboral, porque no se puede pretender que el administrador de la Base de datos sea irremplazable o único, sino que también se pueden producir procesos transparentes de auditoría de sistemas, comportamiento, estructura y administración de las Bases de datos.

Cuando se emplea el modelo E-R se utilizan los siguientes elementos gráficos:

- Entidad 
- Atributos 
- Relación 
- Conector 

Además, es importante tener en cuenta los siguientes aspectos para el diseño del diagrama E-R.

1. Una entidad se relaciona con otra entidad con una línea continua, ya que no lleva flechas.
2. Toda relación debe de llevar una cardinalidad.
3. Una relación siempre se hace entre dos entidades por medio de un rombo. Por ejemplo, si se tiene una entidad **alumno** y otra entidad **asignatura**, se traza una línea. Y en medio de la línea se ubica un rombo, en donde se escribe: "el alumno se



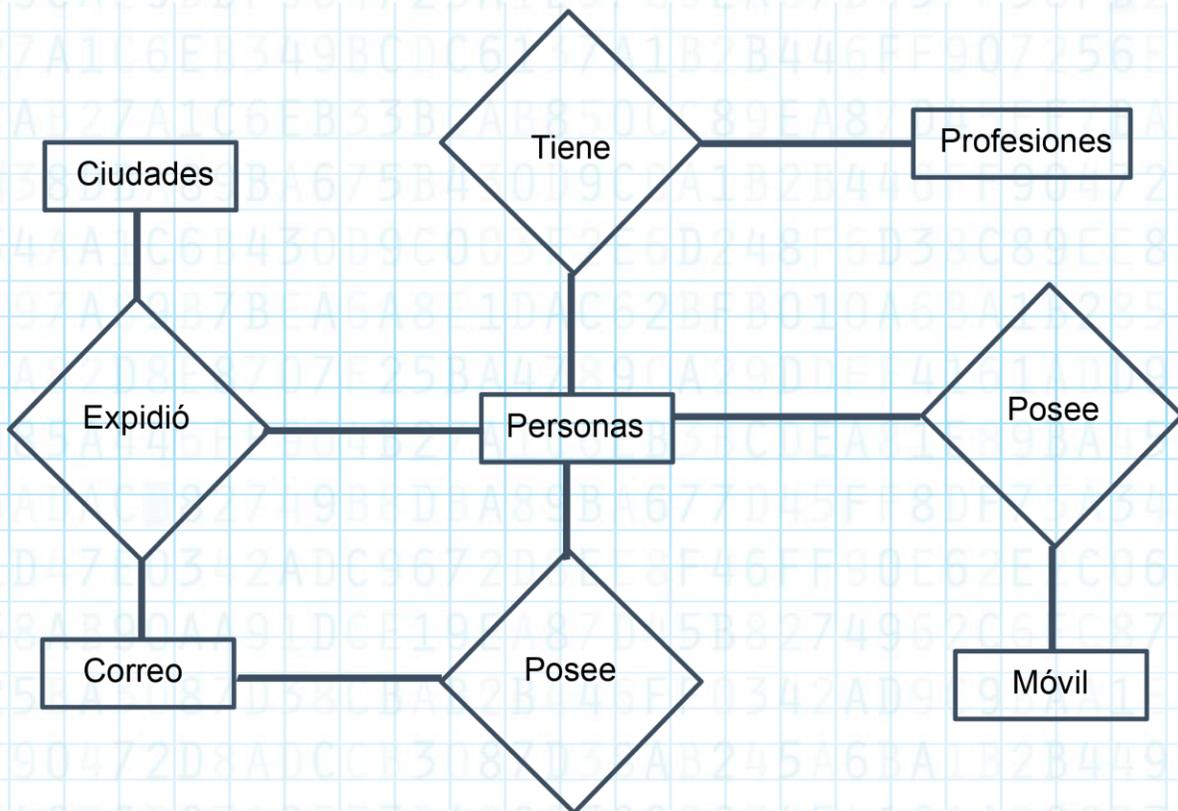
NORMALIZACIÓN Y EL MODELO ENTIDAD RELACIÓN (E-R)

inscribe". Entonces el nivel sería uno a muchos, ya que el alumno se inscribe a varias asignaturas.



4. Cada entidad deberá tener sus atributos o campos.

Entonces, tomando como ejemplo el resultado de la normalización, a continuación se presenta el diagrama E-R.



GLOSARIO

Atributo: Es un elemento susceptible de tomar valores y se denomina a cada columna de la entidad o tabla.

Entidad: Objetos principales en donde se almacena información.

Relación: Asociación entre entidades.

Valor Atómico: Es un solo valor por cada uno de los campos que conforman a los atributos de las entidades.

Valor Nulo: Representa un valor para un atributo que es actualmente desconocido o no es aplicable para ese registro.

Valor Repetido: Representa un valor para un atributo que puede estar dos o más veces.



BIBLIOGRAFÍA

C.J. Date, S. y Ruiz, M. (2001). *Introducción a los sistemas de bases de datos*. (Vol. 7). México: Prentice Hall.

Henry F. K, Silberschatz, A. y Sudarshan, S. (2006). *Fundamentos de bases de datos* (Vol. 5). Madrid, España: McGraw-Hill.



NORMALIZACIÓN Y EL MODELO ENTIDAD RELACIÓN (E-R)

WEBGRAFÍA

García, Roberto. (2009) Apuntes de cómo normalizar una tabla de bases de datos. http://roberto-garciac.blogspot.com/2009/12/apuntes-de-como-normalizar-una-tabla-de_16.html, consultado en julio de 2014.

Wikipedia. (2014). *Normalización de Bases de Datos*. http://es.wikipedia.org/wiki/Normalizaci%C3%B3n_de_bases_de_datos, consultado en abril de 2014.

Soporte Microsoft. (2014). *Fundamentos de la Normalización de bases de datos*. <http://support.microsoft.com/kb/283878/es>, consultado en julio de 2014.

Mysql hispano. (2003). *Normalización de Bases de Datos*. <http://www.eet2mdp.edu.ar/alumnos/MATERIAL/MATERIAL/info/infonorma.pdf>, consultado en julio de 2014.



NORMALIZACIÓN Y EL MODELO ENTIDAD RELACIÓN (E-R)

CRÉDITOS

Líderes del Proyecto

Carlos Fernando Cometa Hortúa
Juan Pablo Vale Echeverry

Líder de Línea

Julián Andrés Mora Gómez

Experta Temática

Carmen Patricia del Rocío Lozano Ramírez

Asesora Pedagógica

Yiced Pulido Cabezas

Editora

Luisa Fernanda Boada Téllez

Equipo de Diseño

Lina Marcela García López
Dalys Ortega Caicedo
Nazly María Victoria Díaz Vera
Yisell Noraya Buitrago Rodríguez
Liam Pamela Hermida García

Equipo de Programación

Luis Fernando Amórtegui García
Charles Richar Torres Moreno
Carlos Andrés Orjuela Lasso
Sergio Alejandro Camacho Hernández
Edilson Laverde Molina

