

Materiales complementarios

5. Normalización



Contenidos:	Epígrafes unidades:
1. Diagrama entidad/relación extendido	Unidad 3 (1.,2.,3.,4.,5.,6.,7.,8.)
2. Modelo relacional	Unidad 2 (1.,2.,3.,4.,5.)
3. Transformación MER-Modelo relacional	Unidad 3 (9.,10.)
4. Normalización	Unidad 3 (11.)

Normalización

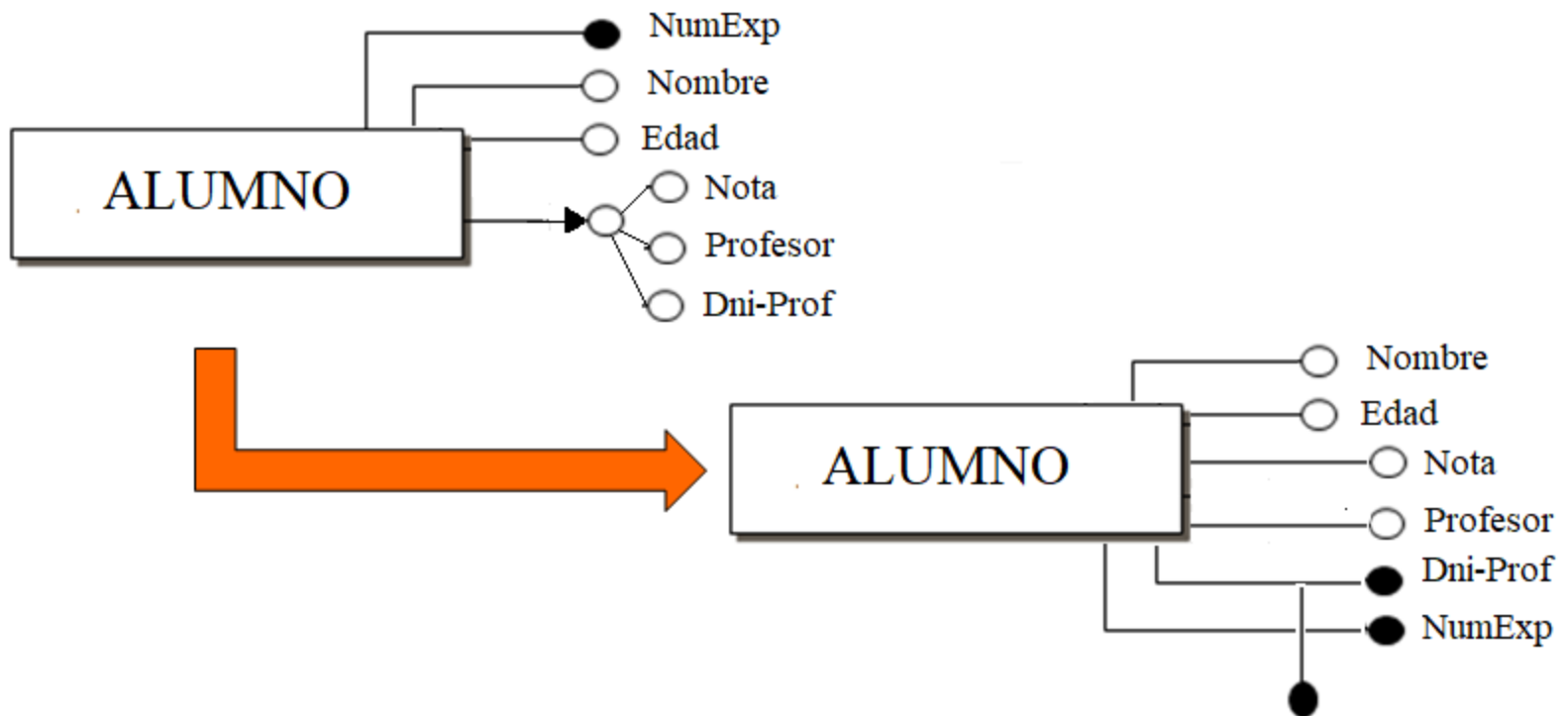


Qué es normalizar

Eliminar la REDUNDANCIA (repetición) y las INCONSISTENCIAS de dependencia en el DISEÑO de las tablas.

Normalización

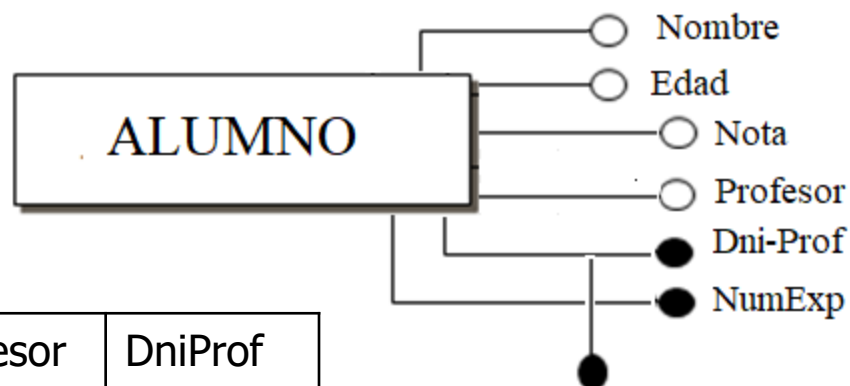
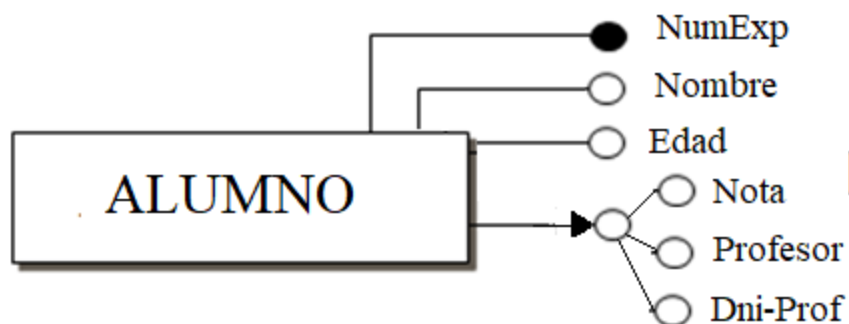
Ejemplo de diseño inadecuado:



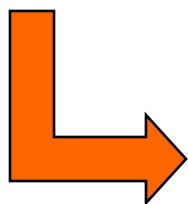
Si transformamos el atributo multivaluado (porque no se permiten en el modelo relacional) de esta manera el resultado es una entidad en la que para un alumno que tenga más de un profesor se van a repetir sus datos (Nombre, Edad y NumExp en varias ocurrencias: una por cada profesor que tenga)

Normalización

Ejemplo de diseño inadecuado:



NumExp	Nombre	Edad	Nota	Profesor	DniProf
1111	Luis	18	7	Juan M.	99999
			8	Ana G.	44444
2222	Marta	20	5	Juan M.	99999
.....



NumExp	Nombre	Edad	Nota	Profesor	DniProf
1111	Luis	18	7	Juan M.	99999
1111	Luis	18	8	Ana G.	44444
2222	Marta	20	5	Juan M.	99999
.....

Normalización

Ejemplo de diseño inadecuado

NumExp	Nombre	Edad	Nota	Profesor	DniProf
1111	Luis	18	7	Juan M.	99999
1111	Luis	18	8	Ana G.	44444
2222	Marta	20	5	Juan M.	99999
.....

❑ ¿Por qué?

- Redundancias
- Necesidad de actualización en cadena
- Posibles inconsistencias en las actualizaciones
- Necesidad de valores nulos



NO PODEMOS DEJARLO ASÍ



NORMALIZAR

Normalización

Por qué normalizar:

Prevenir problemas y anomalías al insertar, modificar o borrar tuplas

Ejemplo, esta tabla contiene los datos de alumnos y sus asignaturas:

**ALUMNO (MUMEXP, DNIPROF, NOMBRE-ALUM, EDAD-ALUM
NOMBRE-PROF, NOTA)**

- ❑ Al añadir alumno → añadir profesor (son clave primaria)**
- ❑ Modificar un nombre de profesor, o edad de un alumno → en todas las tuplas en las que aparezcan**
- ❑ Al borrar un alumno → puede perderse toda la información de un profesor (si no lo tiene ningún otro alumno)**

Normalización



En qué consiste

Descomposición de la relación universal (tabla que contiene todos los atributos) o de una colección de relaciones equivalentes a la misma, en una colección de relaciones en la que las anomalías de actualización (inserción, borrado y modificación) no existan o sean mínimas.

- ❑ Ir descomponiendo relaciones en otras de menor tamaño (grado) de forma que satisfaga una serie de restricciones específicas (Formas Normales)**
- ❑ Esta técnica se emplea en el diseño conceptual (modelo E-R) y/o en el diseño lógico de base de datos.**

Normalización

Objetivos

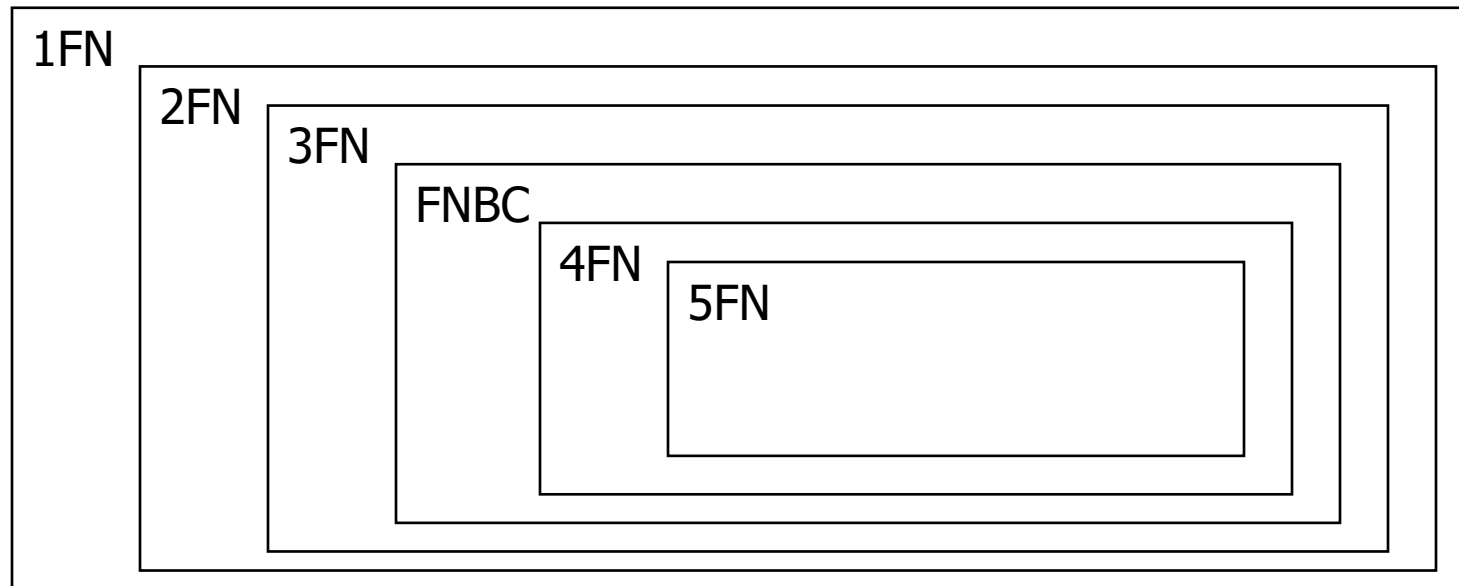
- ❑ Eliminar anomalías de actualización
- ❑ Conservar la información (descomposición sin pérdida de información)
- ❑ Conservar las dependencias funcionales (descomposición sin pérdida de DF)
- ❑ No crear dependencias nuevas o interrelaciones inexistentes
- ❑ Facilidad de uso
- ❑ Eficiencia

¡A VECES NO SON COMPATIBLES!

Normalización

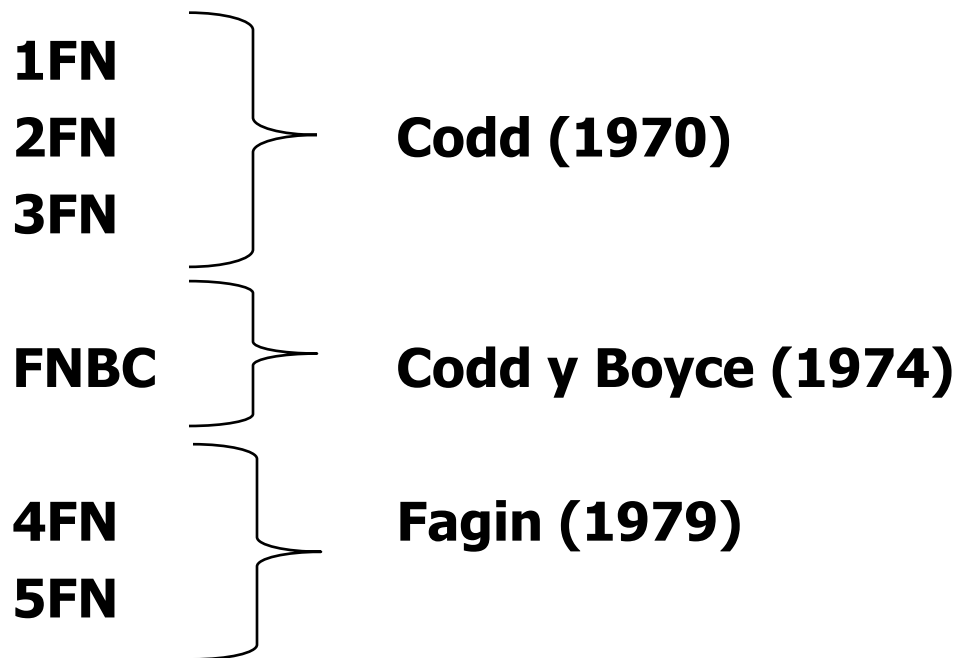
Existen 6 Formas Normales

Una relación está en una forma normal particular si satisface la anterior y cierto conjunto específico de restricciones particulares



Normalización

Formas normales



Normalización



Dependencias (ver Unidad 3 apartado 11.)

Normalización

El problema

Crear una base de datos que almacene la información de los usuarios de la empresa E-COMMERCE LTDA que permita comunicarles novedades de los productos, promociones, enviar las facturas, tanto por correo tradicional como electrónico.

Normalización

Listado de Información

- ☐ **Nombre del cliente**
- ☐ **Empresa donde trabaja**
- ☐ **Dirección empresa**
- ☐ **E-mail (todos los que pueda tener)**
- ☐ **Dirección residencia**
- ☐ **Teléfono**

Normalización

Tabla sin NINGUNA regla de normalización:

	Tabla usuarios				
dni	nombre	empresa	Dir_empresa	email1	email2
111	Oscar Martínez	USabana	Pte comun Chia	omc@usabana.edu.co	oqui@hotline.com
222	Yadira Rincón	Digital Co	Aut Norte con 127	yaya@usabana.edu.co	yayis@latinmail.com

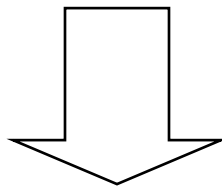
¿Qué hacer si necesitamos colocar otra e-mail?

Normalización

¿Creamos un campo nuevo email3?

¡No siempre es viable!

Debemos crear un sistema de base de datos funcional que pueda crecer y adaptarse a los nuevos requisitos.



Base de datos normalizada

Normalización



PRIMERA FORMA NORMAL

Una relación está en 1FN cuando cada atributo toma un solo valor del dominio.

- **Eliminar los grupos repetitivos (atributos multivaluados) de las tablas.**
- **Identificar cada grupo de datos relacionados con una clave primaria.**
- **Por definición, el modelo relacional no admite atributos multivaluados por lo que al hacer las transformaciones previas del MER las tablas deben quedar en 1FN**

Normalización

La tabla en primera forma normal

usuarios				
dni	nombre	empresa	Direccion empresa	Email
111	Oscar Martínez	ABC	Usabana	omc@usabana.edu.co
111	Oscar Martínez	ABC	Usabana	oqui@hotline.com
222	Yadira Rincon	XYZ	Digital co	yaya@usabana.edu.co
222	Yadira Rincon	XYZ	Digital co	yayis@latinmail.com

Definimos clave primaria compuesta

Hemos eliminado grupos repetitivos

SEGUNDA FORMA NORMAL

Una relación está en 2FN si y sólo si está en 1FN y cada **atributo no principal** tiene una dependencia funcional completa respecto de alguna de las claves

- **Crear tablas separadas para aquellos grupos de datos que no tengan dependencia funcional completa**
- **Relacionar estas tablas mediante una clave ajena.**

Atributo principal: el que forma parte de alguna clave candidata

Normalización

La tabla en primera forma normal

usuarios				
dni	nombre	empresa	Direccion empresa	Email
111	Oscar Martínez	ABC	Usabana	omc@usabana.edu.co
111	Oscar Martínez	ABC	Usabana	oqui@hotline.com
222	Yadira Rincon	XYZ	Digital co	yaya@usabana.edu.co
222	Yadira Rincon	XYZ	Digital co	yayis@latinmail.com

clave primaria
compuesta

Grupos de datos que
dependen sólo de dni

Normalización

La tabla en segunda forma normal

usuarios			
dni	nombre	empresa	direccion_empresa
111	Oscar Martínez	Usabana	Puente comun Chía
222	Yadira Rincón	Digital co	Aut Norte con 127

Grupo de datos que dependen de dni

Relacionamos las tablas

emails	
dni	email
111	omc@usabana.edu.co
111	oqui@hotline.com
222	yaya@usabana.edu.co
222	yayis@latinmail.com

Normalización



TERCERA FORMA NORMAL

Una relación está en 3FN si y sólo si está en 2FN y no existe ningún atributo no principal que dependa transitivamente de la clave primaria

- **Eliminar aquellos campos que no dependan de la clave primaria.**
- **Crear tablas separadas para aquellos grupos de datos que tengan dependencias transitivas**

Normalización

La tabla en segunda forma normal

usuarios			
dni	nombre	empresa	direccion_empresa
111	Oscar Martínez	Usabana	Puente comun Chía
222	Yadira Rincón	Digital co	Aut Norte con 127

emails	
dni	email
111	omc@usabana.edu.co
111	oqui@hotline.com
222	yaya@usabana.edu.co
222	yayis@latinmail.com

La dirección de la empresa no dependen del id del usuario, sino de empresa

Normalización

La tabla en tercera forma normal

usuarios		
dni	nombre	empresa
111	Oscar Martínez	Usabana
222	Yadira Rincón	Digital co

empresas	
empresa	direccion_empresa
Usabana	Puente comun Chía
Digital co	Aut Norte con 127

emails	
dni	email
111	omc@usabana.edu.co
111	oqui@hotline.com
222	yaya@usabana.edu.co
222	yayis@latinmail.com

Se establecen
las relaciones

Normalización

Las tablas en tercera forma normal

usuarios		
dni	Nombre	Id empresa
111	Oscar Martínez	1
222	Yadira Rincón	2

Se establecen
las relaciones

empresas		
Id empresa	empresa	direccion_empresa
1	Usabana	Puente comun Chia
2	Digital co	Aut Norte con 127

Los datos de la
empresa
dependen de la
Id. de la
empresa

emails	
dni	Email
111	omc@usabana.edu.co
111	oqui@hotline.com
222	yaya@usabana.edu.co
14209	yayis@latinmail.com

Normalización

FORMA NORMAL BOYCE-CODD (FNBC)

Una relación R está en FNBC si y sólo si está en 3FN y todo determinante es clave candidata

- Una R en FNBC también está en 3FN
- Si en una relación existen dependencias funcionales de algún conjunto de atributos que no sean una clave candidata, se puede descomponer sin pérdidas en un conjunto de relaciones equivalentes en FNBC.
- Contempla el caso en el que una relación tiene dos o más claves candidatas compuestas y solapadas (comparten algún atributo)

- R(dni, exp, asig, nota)
dni, asig \rightarrow nota
exp, asig \rightarrow nota
exp \rightarrow dni
dni \rightarrow exp

Claves candidatas: dni, asig y exp, asig

exp y dni no son clave candidata

Uno de ellos debe estar en otra relación