

## 8. Modelos de datos

### 8.1. Introducción

Un **modelo de datos** es un grupo de herramientas conceptuales que nos permiten describir los datos, sus relaciones, su semántica y sus limitaciones, de tal forma que facilita la interpretación de nuestro mundo real y su representación en forma de datos en un sistema informático.

#### 8.1.1. Clasificación

Los modelos de datos se aplican a tres niveles: externo, conceptual e interno. Los correspondientes a los dos primeros son los **modelos de datos lógicos** y los correspondientes al último los **modelos de datos físicos**.

Dentro del grupo de los modelos lógicos, podemos diferenciar:

- Los **modelos lógicos basados en objetos**, en cuyo caso el elemento de referencia es el objeto, entendido como tal aquel que existe y que puede distinguirse de otros. Son modelos abstractos y muy flexibles.

El más conocido de estos modelos es el **Modelo Conceptual** o **Modelo Entidad–Relación**.

- Los **modelos lógicos basados en registros**, en cuyo caso el elemento básico es una ocurrencia o conjunto de datos relacionados de algún modo. Son modelos más concretos y más cercanos a la implementación.

Los más conocidos son:

- Modelo Jerárquico
- Modelo en Red
- Modelo Relacional

### 8.2. Modelo Conceptual de Datos

El Modelo Conceptual de Datos, también llamado Modelo Entidad–Relación, es una técnica especial de representación gráfica que incorpora información referente a los

## 8. Modelos de datos

datos y las relaciones existentes entre ellos, para darnos una visión del mundo real.

Es un paso previo al futuro diseño de las bases de datos del sistema.

Las características del Modelo Conceptual son:

1. Reflejan tan sólo la existencia de los datos, no lo que se hace con ellos.
2. Se incluyen todos los datos del sistema en estudio, y por tanto no está orientado a aplicaciones particulares.
3. Es independiente de las bases de datos y sistemas operativos concretos que puedan utilizarse.
4. No se tienen en cuenta restricciones de espacio, almacenamiento, ni tiempo de ejecución.
5. Está abierto a la evolución del sistema.

Por tanto, en el Modelo Conceptual se da una visión del mundo real con la mayor naturalidad, mediante los datos y sus relaciones, de tal manera que pueda implementarse dicha visión en una base de datos que mantenga sus propiedades.

El Modelo Conceptual, en definitiva, se basa en la percepción de un mundo real que consiste en un conjunto de objetos básicos denominados Entidades, así como las Relaciones existentes entre ellos.

### 8.2.1. Conceptos fundamentales del Modelo Conceptual

#### 8.2.1.1. Entidad

Una **entidad** es una cosa u objeto concreto o abstracto que existe, que puede distinguirse de otros y del cual se desea almacenar información.

#### 8.2.1.2. Instancias de entidades

Las entidades se refieren colectivamente a un conjunto de casos concretos de entidades, todas del mismo tipo. A cada caso concreto de una entidad se la denomina **instancia (de una entidad)**.

Cada instancia de una entidad debe ser distinta de las demás, es decir, debe poder distinguirse de las otras.

#### Ejemplo

La entidad *Cliente* podría tener como instancia suya a «*Juan Pérez Martín, con DNI nº 46.347.187-B, y que vive en la C/. Larga, nº 15*».

### 8.2.1.3. Tipos de entidades

Entre las entidades se pueden diferenciar dos tipos:

1. **Entidades fuertes:** son aquellas cuyas instancias son identificables y existen por sí mismas.

*Ejemplo:* la entidad Cliente.

2. **Entidades débiles:** son aquellas cuyas instancias son identificables y existen solamente por estar asociadas a otra u otras entidades.

*Ejemplo:* la entidad Línea de Factura. Una línea de factura sólo existe como parte de una factura, y no tiene sentido por sí sola.

### 8.2.1.4. Atributo

Un **atributo** es una unidad básica e indivisible de información acerca de una entidad, que sirve para identificarla o describirla.

Los atributos son, por tanto, los elementos simples de datos que componen una entidad.

Todas las instancias de una entidad tienen los mismos atributos, aunque en cada caso esos atributos tendrán valores distintos.

#### Ejemplo

Los atributos de la entidad Cliente, podrían ser DNI, nombre, dirección, población, provincia, código postal, teléfono, etc.

### 8.2.1.5. Superclave

Es un conjunto de uno o más atributos que permiten identificar de forma única a una instancia dentro de una entidad.

#### Ejemplo

En la entidad Cliente, podrían ser superclaves:

- DNI + nombre + apellidos
- DNI
- DNI + teléfono
- nombre + apellidos + teléfono

## 8. Modelos de datos

### 8.2.1.6. Claves candidatas

Son aquellas superclaves que son mínimas, es decir, que ningún subconjunto suyo es a su vez superclave.

#### Ejemplo

De las superclaves anteriores, serían claves candidatas:

- DNI
- nombre + apellidos + teléfono

### 8.2.1.7. Clave primaria y claves alternativas

De entre todas las claves candidatas, el analista elegirá la más apropiada (en base a su tamaño, sencillez, facilidad de utilización, etc.), que pasará a denominarse **clave primaria**. El resto de claves candidatas pasarán a denominarse **claves alternativas** (o **secundarias**).

#### Ejemplo

De las claves candidatas anteriores, la más apropiada, por ser la más corta y la más sencilla, es el DNI. Por tanto, la clave primaria de la entidad Cliente estaría formada por un único atributo llamado DNI. El conjunto de atributos «nombre + apellidos + teléfono» sería una clave alternativa.

### 8.2.1.8. Relación

Una **relación** es una asociación, sin existencia propia, de varias entidades.

#### Ejemplo

Entre las entidades Alumno y Asignatura, se establece la relación «Estudia». De manera que podemos decir que «los alumnos *estudian* asignaturas».

Las relaciones son **bidireccionales**, por tanto también podemos decir que «las asignaturas *son estudiadas* por los alumnos».

### Ejemplo

Entre las entidades Cliente y Factura se establece la relación «es facturada a», de manera que podemos decir que «las facturas *son facturadas a* los clientes». En el otro sentido, podríamos decir (aunque no quede muy bien) que «los clientes *generan* facturas (al comprar algo)».

#### 8.2.1.9. Atributos de relaciones

Los **atributos de una relación** son:

- aquellos atributos que forman parte de las entidades que asocia, y
- (posiblemente) otros atributos que formen parte de la propia relación, y no de ninguna entidad que asocie.

### Ejemplo

Los atributos de la relación «Estudia» serían los propios de la entidad Alumno, más los propios de la entidad Asignatura. En este caso, la relación no posee ningún atributo propio, sino que sus atributos son los de las entidades que asocia.

### Ejemplo

En un sistema de gestión de video-clubs, supongamos la relación «alquila», entre las entidades Socio y Película (es decir, que «los clientes *alquilan* películas»). Los atributos de dicha relación serían los propios de las entidades Socio y Película, más el atributo «Fecha de alquiler», que indicaría la fecha en la que el socio se llevó la película, más el atributo «Fecha de devolución», que indicaría la fecha en la que el socio entregó esa película. Estos dos atributos no forman parte de la entidad Socio ni de la Entidad película. Son atributos propios de la relación.

#### 8.2.1.10. Instancias de relaciones

Las relaciones se refieren colectivamente a un conjunto de casos concretos de relaciones, todas del mismo tipo. A cada caso concreto de una relación se la denomina **instancia (de una relación)**.

## 8. Modelos de datos

### Ejemplo

Instancias concretas de la relación *estudia*, existente entre las entidades Alumno y Asignatura, podrían ser:

- Juan estudia Programación.
- Juan estudia Sistemas.
- Antonio estudia Matemáticas.
- María estudia Inglés.
- Rosa estudia Sistemas.

Como podemos ver, una instancia de una entidad puede estar relacionada al mismo tiempo con varias instancias de la otra entidad, y viceversa. Cada una de esas asociaciones es una instancia de la relación.

#### 8.2.1.11. Dimensión de una relación

Es el número de entidades que participan en ella. Las relaciones más comunes son las *binarias*, que son aquellas en las que intervienen sólo dos relaciones. También existen relaciones *ternarias*, pero son muchísimo menos comunes.

#### 8.2.1.12. Clave de una relación

La **clave de una relación** identifica de forma única a una instancia concreta de una relación, de entre todas las instancias posibles. Se forma mediante la concatenación de las claves primarias de las entidades que asocia, y en ocasiones algunos atributos propios de la relación.

### Ejemplo

Supongamos la relación «son facturadas a», entre las entidades Cliente y Factura. La clave primaria de la entidad Cliente es el DNI. La clave primaria de la entidad Factura es el N° de factura. Por tanto, la clave primaria de la relación sería «DNI + N° de factura».

### Ejemplo

Supongamos la relación «estudia», entre las entidades Alumno y Asignatura. La clave primaria de la entidad Alumno es el DNI. La clave primaria de la entidad Asignatura es el Código de la asignatura. Por tanto, la clave primaria de la relación sería «DNI + Código de la asignatura».

### Ejemplo

Supongamos la relación «alquila», entre las entidades Socio y Película. La clave primaria de la entidad Socio es el DNI (o el n° de socio). La clave primaria de la entidad Película es el Código de la película (también podría ser su n° de expediente, por ej.). Sin embargo, la clave de la relación «alquila» no puede ser simplemente «DNI + Código de la película», ya que un socio puede alquilar la misma película varias veces (en días distintos). Por tanto, para identificar de forma única una instancia concreta de dicha relación, necesitamos además el atributo “Fecha de alquiler”, y por tanto la clave de la relación estaría formada por los atributos «DNI + Código de película + Fecha de Alquiler».

Esto se ve así:

- el cliente 35.236.264-J alquila la película 38947 el día 11/02/99, y la devuelve el 12/02/99.
- el cliente 46.931.763-P alquila la película 74857 el día 24/03/99 y la devuelve el 25/03/99.
- el cliente 75.486.365-T alquila la película 38947 el día 05/06/99 y la devuelve el 07/06/99.
- el cliente 75.486.365-T alquila la película 14180 el día 05/06/99 y la devuelve el 07/06/99.
- el cliente 35.236.264-J alquila la película 38947 el día 17/08/99 y la devuelve el 18/08/99.

Si tomáramos como clave «DNI + Código de la película», no podríamos distinguir las instancias primera y última, ya que ambas tienen el mismo DNI y el mismo código de película, diferenciándose sólo en la fecha.

### 8.2.2. Representación del Modelo Conceptual

El Modelo Conceptual de Datos se representa mediante un diagrama llamado **diagrama Entidad-Relación**, o simplemente, **diagrama E-R**. Por ello, al Modelo Conceptual a veces se le denomina *Modelo E-R*.

Los diferentes elementos estudiados hasta ahora se representan en un diagrama E-R de la siguiente forma:

1. *Entidad fuerte*: rectángulo con nombre dentro.
2. *Entidad débil*: dos rectángulos concéntricos con nombre dentro.
3. *Relación*: rombo con nombre dentro.
4. *Atributo*: elipse con nombre dentro, o círculo pequeño con nombre al lado.
5. *Clave primaria*: el nombre del atributo subrayado, o el círculo pequeño relleno.

## 8. Modelos de datos

En un diagrama E-R, no se representan las instancias.

### 8.2.3. Otros conceptos relacionados con las Relaciones

#### 8.2.3.1. Grado de una relación

Representa la participación en la relación de cada una de las entidades afectadas, y existen tres tipos posibles. Estos son:

- **una a una (1:1)**: a cada instancia de una entidad le corresponde no más de una instancia de la otra, y viceversa.
- **una a muchas (1:N)**: a cada instancia de la primera entidad pueden corresponderle varias instancias de la segunda, y a cada instancia de la segunda le corresponde no más de una instancia de la primera.
- **muchas a muchas (M:N)**: a cada instancia de la primera entidad pueden corresponderle más de una instancia de la segunda, y viceversa.

El grado de una relación se puede representar en el diagrama E-R, pero es más descriptivo utilizar la cardinalidad.

#### Ejemplo

La relación «Estudia», entre las entidades Alumno y Asignatura, es una relación de muchos a muchos, porque un alumno puede estudiar muchas asignaturas, y una misma asignatura puede ser estudiada por muchos alumnos.

#### Ejemplo

La relación «es facturada a», entre las entidades Cliente y Factura, es una relación de uno a muchos, ya que un cliente puede tener más de una factura de una misma empresa (porque le haya comprado varios artículos en varios momentos distintos), pero una factura sólo pertenece a un cliente.

#### 8.2.3.2. Cardinalidad de una entidad

La **cardinalidad de una entidad en una relación** representa el mínimo y el máximo de la participación de dicha entidad en la relación. Es decir: representa el número mínimo y máximo de instancias de esa entidad que pueden asociarse con *una* instancia de la otra entidad en la relación. Existen las siguientes posibilidades:

- **(1,1)**: indica que a cada instancia de una entidad le corresponde otro en la otra entidad (obligatoriedad).



- **(0,1)**: indica que se pueden corresponder una o ninguna instancia de una entidad con una instancia de la otra entidad (opcionalidad).
- **(1,n)**: indica que se pueden corresponder una o más instancias de una entidad con una instancia de la otra entidad (obligatoriedad).
- **(0,n)**: indica que se pueden corresponder ninguna, una o más instancias de una entidad con una instancia de la otra entidad (opcionalidad).

La cardinalidad máxima de una entidad en una relación coincide con el grado de la relación para esa entidad. Además, la cardinalidad se representa en el diagrama E-R, junto a la entidad correspondiente.

### Ejemplo

En la relación «Estudia», la entidad Alumno tiene cardinalidad (0,n) (porque una asignatura puede ser estudiada por cero, uno o más alumnos), y la entidad Asignatura tiene cardinalidad (1,n) (porque un alumno puede tener una o más asignaturas, al menos una).

### Ejemplo

En la relación «es facturada a», la entidad Cliente tiene cardinalidad (1,1) (porque una factura es de un cliente, y sólo de uno, y cada factura debe ir a un cliente; no puede haber una factura sin cliente), y la entidad Factura tiene cardinalidad (0,n) (porque un cliente puede tener cero, una o más facturas de la misma empresa).

## 8.2.4. Conversión del Modelo Conceptual al Modelo Relacional

El objetivo es obtener un conjunto de tablas interrelacionadas entre sí, a partir de las entidades y las relaciones identificadas en el Modelo Conceptual. Las tablas se relacionan por medio de sus claves primarias, como se verá más adelante.

Para ello, se tendrán en cuenta las siguientes reglas:

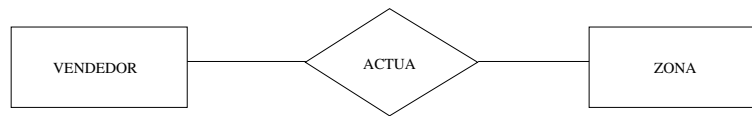
### Transformación de entidades

Una entidad se transforma en una tabla. Los atributos de la entidad serán atributos (o columnas) de la tabla. La clave primaria de la tabla será la de la entidad.

## 8. Modelos de datos

### Transformación de relaciones

Se apoyará en el siguiente ejemplo:



Para la entidad VENDEDOR, tendremos los atributos CODVEN, NOMVEN y TELVEN. Para la entidad ZONA, tendremos dos atributos: CODZON y NOMZON.

Las tablas con sus instancias son las siguientes:

<u>CODVEN</u>	NOMVEN	TELVEN
016	López	4421311
121	García	5070010
208	Martín	2030887
152	Sánchez	3521414

<u>CODZON</u>	NOMZON
1	Vallecas
2	Centro
3	Aluche
4	Argüelles
5	Batán
6	Retiro

Se dan los siguientes casos:

1. Dependiendo de las *cardinalidades de las entidades relacionadas*, tenemos:

- a) **(1,1) ↔ (1,1)**: Sólo se necesita una tabla, que tendrá como atributos los de las dos entidades asociadas. La clave primaria de la tabla puede ser cualquiera de las claves primarias de cada entidad, pero se tomará la que pertenezca a la entidad más dominante.

En nuestro ejemplo:

<u>CODVEN</u>	NOMVEN	TELVEN	<u>CODZON</u>	NOMZON
016	López	4421311	2	Centro
121	García	5070010	4	Argüelles
208	Martín	2030887	1	Vallecas
152	Sánchez	3521414	3	Aluche

## 8.2. Modelo Conceptual de Datos

Las filas correspondientes a las zonas Batán y Retiro no podrían existir simplemente porque no están asociadas a ningún vendedor.

- b)  $(1,1) \leftrightarrow (0,1)$ : Se usa una única tabla que mantenga a las dos entidades, cuya clave será la de la entidad obligatoria (en algunos casos, la otra tendrá valores nulos).

En nuestro ejemplo, supongamos que ZONA fuese obligatoria pero no VENDEDOR:

<u>CODVEN</u>	<u>NOMVEN</u>	<u>TELVEN</u>	<u>CODZON</u>	<u>NOMZON</u>
016	López	4421311	1	Vallecas
121	García	5070010	3	Aluche
208	Martín	2030887	2	Centro
152	Sánchez	3521414	4	Argüelles
			5	Batán
			6	Retiro

Si se quiere evitar los valores nulos, una alternativa sería descomponer la tabla en dos, añadiendo a la no obligatoria la clave primaria de la entidad obligatoria:

<u>CODVEN</u>	<u>NOMVEN</u>	<u>TELVEN</u>	<u>CODZON</u>
016	López	4421311	1
121	García	5070010	3
208	Martín	2030887	2
152	Sánchez	3521414	4

<u>CODZON</u>	<u>NOMZON</u>
1	Vallecas
3	Aluche
2	Centro
4	Argüelles
5	Batán
6	Retiro

- c)  $(0,1) \leftrightarrow (0,1)$ : Se usan dos tablas, cada una para una entidad, añadiendo en la tabla de la entidad más dominante la clave de la otra entidad (*clave ajena*), que en algunos casos tendrá valores nulos.

En nuestro ejemplo, si utilizáramos una sola tabla, tendríamos:

## 8. Modelos de datos

CODVEN	NOMVEN	TELVEN	CODZON	NOMZON
016	López	4421311	4	Argüelles
			2	Centro
208	Martín	2030887	1	Vallecas
152	Sánchez	3521414		
121	García	5070010	3	Aluche
			5	Batán
			6	Retiro

Vemos que aparecen valores nulos y no es posible determinar la clave.

La solución consiste en usar dos tablas:

CODVEN	NOMVEN	TELVEN	CODZON
016	López	4421311	4
208	Martín	2030887	1
152	Sánchez	3521414	
121	García	5070010	3

CODZON	NOMZON
1	Vallecas
2	Centro
3	Aluche
4	Argüelles
5	Batán
6	Retiro

- d)  $(1,1) \leftrightarrow (0,n)$  ó  $(1,n)$ : Se necesitan dos tablas, una para cada entidad. La clave de cada tabla será la clave de la entidad correspondiente. La clave de la entidad  $(1,1)$  se añade como un atributo más a la tabla de la entidad  $(0,n)$  ó  $(1,n)$ .

En nuestro ejemplo, una sola tabla contendría:

CODVEN	NOMVEN	TELVEN	CODZON	NOMZON
016	López	4421311	1	Vallecas
016	López	4421311	3	Aluche
208	Martín	2030887	5	Batán
152	Sánchez	3521414	2	Centro
152	Sánchez	3521414	4	Argüelles
152	Sánchez	3521414	6	Retiro
121	García	5070010		

Esta tabla presenta dos problemas: aparecen valores nulos en las filas en las que el vendedor no tiene asignada zona, y aparecen valores repetidos cuando un vendedor actúa en más de una zona. Se descompone en dos tablas:

## 8.2. Modelo Conceptual de Datos

<u>CODVEN</u>	NOMVEN	TELVEN
016	López	4421311
208	Martín	2030887
152	Sánchez	3521414
121	García	5070010

<u>CODZON</u>	NOMZON	CODVEN
1	Vallecas	016
2	Centro	152
3	Aluche	016
4	Argüelles	152
5	Batán	208
6	Retiro	152

**Caso particular:** si la entidad con cardinalidad (0,n) ó (1,n) es débil respecto de la otra entidad, la clave de su tabla correspondiente estará formada por la unión de la clave de la entidad débil y la de la entidad fuerte.

- e) **(0,1) ↔ (0,n) ó (1,n):** Se utilizan dos tablas, una con cada entidad. La clave de la entidad (0,1) se añade como un atributo más a la tabla de la entidad (0,n) ó (1,n) (clave ajena que en algunos casos tendrá valores nulos).

En nuestro ejemplo, tenemos:

<u>CODVEN</u>	NOMVEN	TELVEN	<u>CODZON</u>	NOMZON
016	López	4421311	1	Vallecas
			2	Centro
016	López	4421311	3	Aluche
121	García	5070010	4	Argüelles
121	García	5070010	5	Batán
208	Martín	2030887	6	Retiro
152	Sánchez	3521414		

La tabla presenta muchos valores nulos y valores repetidos. No es posible determinar la clave. Se descompone en dos tablas:

<u>CODVEN</u>	NOMVEN	TELVEN
016	López	4421311
121	García	5070010
208	Martín	2030887
152	Sánchez	3521414

## 8. Modelos de datos

<u>CODZON</u>	NOMZON	CODVEN
1	Vallecas	016
2	Centro	
3	Aluche	016
4	Argüelles	121
5	Batán	121
6	Retiro	208

**Caso particular:** puede darse el mismo caso particular del apartado anterior.

- f) **(0,n) ó (1,n) ↔ (0,n) ó (1,n):** Se necesitan tres tablas, una para cada entidad, siendo la clave primaria la correspondiente a la entidad que representa, y una tercera formada por las claves de ambas entidades. Las tablas de las entidades se asocian con la tabla de la relación original a través de una relación 1:N, en la que la tabla de la relación original es la que posee la cardinalidad (0,n) o (1,n).

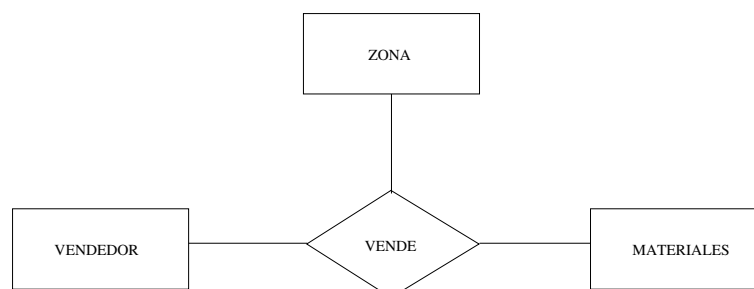
En nuestro ejemplo, tendríamos como solución:

<u>CODVEN</u>	NOMVEN	TELVEN
016	López	4421311
121	García	5070010
208	Martín	2030887
152	Sánchez	3521414

<u>CODZON</u>	NOMZON	<u>CODVEN</u>	<u>CODZON</u>
1	Vallecas	016	1
2	Centro	016	2
3	Aluche	016	3
4	Argüelles	121	2
5	Batán	208	1
6	Retiro	208	3

- g) **Relación ternaria M:N:** Se necesitarán  $(n + 1)$  tablas, una para cada entidad (siendo su clave primaria la clave de la entidad) y una más que contenga las claves de todas las entidades.

Como ejemplo, podemos tener el siguiente:



## 8.2. Modelo Conceptual de Datos

Este diagrama se traduciría en las siguientes tablas:

<u>CODVEN</u>	<u>NOMVEN</u>	<u>TELVEN</u>	<u>CODMAT</u>	<u>NOMMAT</u>
016	López	4421311	A	Alfombras
121	García	5070010	B	Libros
208	Martín	2030887	C	Muebles
152	Sánchez	3521414	D	Electrod.

<u>CODZON</u>	<u>NOMZON</u>
1	Vallecas
2	Centro
3	Aluche
4	Argüelles
5	Batán
6	Retiro

<u>CODVEN</u>	<u>CODZON</u>	<u>CODMAT</u>
016	2	B
208	2	B
208	4	A
152	4	C
152	2	B
121	3	D

2. *Relación con atributos propios (entidad asociativa)*: Considerar los atributos propios de la relación, con el objetivo de incorporarlos a la tabla correspondiente.

*Ejemplo*: La relación «Alquila» entre las entidades «Socio» y «Película» tiene como atributos propios la fecha de alquiler y la fecha de devolución. Dependiendo de las cardinalidades de las entidades, esos atributos formarán parte de la tabla Socio, o de la tabla Película, o quizás de la tabla que represente a la relación Alquila (en el caso de que el grado de la relación fuese «muchos a muchos»).

3. *Entidades débiles*: Tener en cuenta que una entidad débil suele requerir como parte de su clave a la clave primaria de la entidad de la cual depende.

*Ejemplo*: La entidad débil «Línea de factura» incluirá en su clave primaria la clave de la entidad «Cabecera de factura».

Al final, se consigue una representación en la cual:

- Todas las entidades se convierten en tablas, ligadas entre sí.
- Las relaciones «muchos a muchos» se convierten en tablas, ligadas mediante relaciones «uno a muchos».
- Las únicas relaciones «uno a uno» que se mantienen son aquellas en las que las dos tablas relacionadas tienen cardinalidad (0,1).

## 8.3. Ejercicios

### Ejercicio 1

Empresa nacional encargada de la distribución de productos a diferentes mercados.

En las diversas reuniones con los usuarios se han recogido los siguientes requisitos:

- No existen ciudades diferentes con el mismo nombre. Los nombres de los mercados y de los propietarios son únicos.
- Cada mercado codifica los puestos que tiene secuencialmente, empezando todos desde el puesto 001.
- Un propietario puede tener varios puestos, que se pueden encontrar en el mismo mercado o en mercados diferentes. En este caso, generalmente, el nombre del puesto se repite. Se considera que un puesto tiene un único propietario.
- La empresa tiene varios almacenes, y cada almacén distribuye determinada fruta y NO toda. De la fruta que se puede distribuir puede suceder que exista alguna sin stock.
- En los puestos de los mercados puede darse el caso de que determinada fruta no disponga de stock.
- Un almacén distribuye fruta a todos los puestos de determinados mercados. Estos mercados que se asignan a un almacén son aquellos que están ubicados en el ámbito geográfico del almacén. Pero no distribuye a cualquier mercado que le hace un pedido, puesto que cada mercado tiene sus almacenes asignados.
- Una fruta determinada se puede distribuir en diferentes unidades de medida. Por ejemplo, las manzanas se pueden entregar en cajas de 25 y de 30 kilos.
- Una misma fruta se distribuye al mismo precio desde un almacén, independientemente de la cantidad que se pide y del puesto o mercado al que se le vende. Cada fruta posiblemente se distribuya a precios diferentes según el almacén.
- Cada puesto vende la fruta a un precio, que generalmente difiere de un puesto a otro.
- Cada almacén enumera sus facturas secuencialmente cada año. Los dos primeros caracteres son el año, los seis siguientes es un número secuencial que pone cada almacén. Esto indica que dos almacenes en un año van dando números de facturas iguales, se diferencian teniendo en cuenta el código del almacén.
- En una factura deben aparecer los siguientes datos:
  - N° de factura, y fecha.
  - Mercado: código, razón social, dirección, población, provincia, código postal.
  - Almacén: código, CIF, razón social.



- Puesto: código, CIF, nombre propietario.
- Detalle: fruta (código y denominación), unidad de medida, cantidad, precio unitario y total.
- Pie de factura: base imponible, IVA y total factura.

## Ejercicio 2

Construir un diagrama E-R para una compañía de seguros automovilísticos que cuenta con un conjunto de clientes, cada uno de los cuales posee cierto número de vehículos. Cada automóvil está relacionado con un número de accidentes registrados.

## Ejercicio 3

Dibujar los diagramas E-R para cada uno de los siguientes casos:

1. Una empresa de albañilería desea conocer en todo momento la casa en la que están trabajando sus empleados. Se sabe que en cada casa pueden haber varios empleados y no todos tienen por qué estar ocupados en cada instante.
2. Un colegio desea saber las asignaturas que imparte cada profesor. El jefe de estudios ha dispuesto que cada profesor puede dar más de una asignatura y una asignatura puede ser impartida por varios profesores.
3. Ídem al apartado 2, pero añadiendo lo siguiente: «Los alumnos del colegio tienen varias asignaturas por curso, pero no pueden asistir a más de una asignatura al mismo tiempo».

## Ejercicio 4

En un hospital, normalmente hay un número de pacientes que están esperando para ser operados, o bien, en periodo post-operatorio. Dibujar el diagrama E-R para esta situación, teniendo en cuenta lo siguiente:

1. Un paciente es operado por un cirujano y no puede sufrir dos operaciones el mismo día.
2. Un cirujano no puede operar a dos pacientes a la vez.
3. Después de la operación sólo se administra un medicamento y los posibles efectos secundarios dependen sólo de la medicación administrada.

## 8. Modelos de datos

### Ejercicio 5

En la empresa «Proyectos, S.A.», los proyectos de investigación tienen todos un presupuesto y están asignados a uno o varios empleados. Cada uno de ellos invierte un tiempo para realizarlo. Por otro lado, también disponemos de los datos principales de los empleados que están asignados a un proyecto, y el departamento al que pertenece el empleado dentro de la empresa.

Construir el diagrama E-R que represente la información anterior.

### Ejercicio 6

En el análisis de un centro educativo, se dispone de la siguiente información:

Código de los estudios, nombre de los estudios, descripción, nivel y número de años; código de la asignatura, nombre de la asignatura, número de horas y características; seminario de profesor, NRP (número de registro personal), DNI, nombre, primer apellido, segundo apellido, dirección, teléfono, titulación y especialidad; DNI del alumno, y nombre, primer apellido, segundo apellido, dirección, teléfono, fecha de nacimiento del alumno; del padre y de la madre se conocen el DNI, nombre, primer apellido, segundo apellido, dirección, teléfono, profesión y estudios.

Se guarda información del expediente del alumno. Para cada uno se posee varias fichas con el mismo número de registro, código de la asignatura, nota de junio, nota de septiembre y año académico. El número de registro puede ser igual para el mismo alumno en distintos años académicos. Además:

- Conocemos que hay profesores que son jefe de seminario (un sólo jefe por seminario, y hay varios seminarios), y el resto son miembros del seminario.
- Varios profesores pueden realizar la tutoría a los alumnos en una hora de visita del alumno, y una hora de visita de padres.
- El profesor imparte la asignatura en un mismo aula, en distintas horas y diferentes días.
- El alumno asiste en un grupo a la impartición de la asignatura por parte del profesor.
- El alumno se matricula en los estudios actuales en una fecha, con un número de matrícula y un curso.
- La asignatura pertenece a unos estudios.
- La asignatura puede ser impartida por varios profesores.
- El profesor puede impartir la misma asignatura a distintos grupos.
- Los padres pueden tener varios alumnos en el mismo centro.

- En padres queremos centralizar los grupos de madre y de padre.

Se pide:

1. Construir el diagrama E-R que represente la información anterior.
2. Obtener el modelo lógico de datos equivalente, representándolo mediante un diagrama de estructura lógica de datos.

## 8.4. Normalización de tablas relacionales

El proceso de normalización consiste en seguir una serie de pasos o normas que tras ser aplicadas, se obtienen los datos agrupados en diferentes tablas, de tal forma que es la estructura óptima para su implementación, gestión y explotación desde diferentes futuras aplicaciones.

Una tabla se dice que está en una forma normal cuando satisface un conjunto de restricciones impuestas por dicha norma.

Su objetivo es obtener el mayor número de tablas posibles, dejando en cada una de ellas los atributos imprescindibles para representar a la entidad (o relación entre entidades) a la que hace referencia la tabla.

La ventaja fundamental de la normalización es la *minimización de la redundancia*.

En la práctica, al convertir el modelo conceptual de datos en el modelo lógico, siguiendo las reglas dadas en un apartado anterior, las tablas resultantes ya suelen estar muy normalizadas. Sólo quedaría comprobar su normalización.

### 8.4.1. Primera forma normal

Una tabla se dice que está en 1FN si y sólo si los valores que componen cada atributo de una fila son atómicos. Es decir, en un atributo sólo deben aparecer valores elementales y únicos.

#### Ejemplo

La siguiente tabla de materiales en una ferretería:

COD-MAT	Descripción	Medidas
039	Tornillo	3,5 - 5 - 7
067	Arandela	2 - 5
461	Broca	2,5 - 3 - 3,5

## 8. Modelos de datos

Esta tabla no se encuentra en 1FN porque en una fila concreta, el atributo Medidas tiene varios valores, en lugar de un único valor.

Para convertirla a 1FN, se procede así:

Se descompone la tabla en dos:

- Una que contiene la clave y los atributos que tienen valores únicos. La clave será la misma que la tabla original. El nombre de la tabla será la misma que la tabla original.

<u>COD-MAT</u>	<u>Descripción</u>
039	Tornillo
067	Arandela
461	Broca

- Otra que contiene la clave y los atributos con valores múltiples, de tal manera que cada valor aparecerá en una fila distinta de la nueva tabla. La clave estará formada por ambos atributos.

<u>COD-MAT</u>	<u>Medida</u>
039	3,5
039	5
039	7
067	2
067	5
461	2,5
461	3
461	3,5

Las tablas resultantes ya se encuentran en 1FN.

De cualquier forma, la mejor solución estaría en tener dos tablas: una con los materiales, otra con las medidas, y una tercera que vincule los materiales con sus medidas correspondientes. Como vemos, al final nos queda lo mismo que si desde el principio hubiésemos aplicado las reglas conocidas a una relación M:N entre materiales y medidas.

### 8.4.2. Dependencia funcional

Se dice que un atributo (o conjunto de atributos) B *depende funcionalmente* del atributo (o conjunto de atributos) A, y se representa como  $A \rightarrow B$ , si y sólo si cada valor de A se corresponde con un único valor de B.

### Ejemplo

Entre los atributos DNI y Nombre existe una dependencia funcional, ya que un DNI concreto se corresponde con un sólo Nombre asociado. Por tanto:

$\text{DNI} \rightarrow \text{Nombre}$

Lo contrario no siempre es cierto: a partir de un Nombre, no se puede deducir un sólo DNI asociado:

$\text{Nombre} \nrightarrow \text{DNI}$

Igualmente ocurre con:

$\text{DNI} \rightarrow \text{Dirección}$

Aunque:

$\text{Dirección} \nrightarrow \text{DNI}$

Para abreviar, podemos escribir:

$\text{DNI} \rightarrow \text{Nombre} \mid \text{Dirección}$

En muchas ocasiones, para determinar un único valor de un atributo, no nos basta con conocer el valor de otro, sino que es necesario encontrar los valores de varios atributos.

### Ejemplo

Si tenemos los atributos DNI, Empresa y Sueldo, y sabemos que una persona puede trabajar en varias empresas. Entre los atributos DNI y Sueldo no existe dependencia funcional, puesto que un empleado puede ganar varios sueldos dependiendo de la empresa en la que trabaje. Por tanto, sí puede decirse que:

$\text{DNI} + \text{Empresa} \rightarrow \text{Sueldo}$

### 8.4.3. Segunda forma normal

Una tabla se dice que está en 2FN si y sólo si cumple dos condiciones:

1. Está en 1FN.
2. Todo atributo no primario (aquellos que no pertenecen a la clave primaria) depende de la clave completa, y no sólo de una parte de ella.

*Ejemplo:* La tabla EMPLEADOS, que contiene los atributos DNI, Nombre, Empresa y Sueldo. La clave es DNI + Empresa, ya que una persona puede trabajar en más de una empresa. Se verifican las siguientes dependencias:

$\text{DNI} \rightarrow \text{Nombre}$

$\text{DNI} + \text{Empresa} \rightarrow \text{Sueldo}$

## 8. Modelos de datos

La tabla se encuentra en 1FN, pero no se encuentra en 2FN, ya que el atributo Nombre depende de *parte* de la clave, y no de *toda* la clave.

<u>DNI</u>	<u>Empresa</u>	Nombre	Sueldo
52537462	Microsoft	Pepe Pérez	200.000
52537462	IBM	Pepe Pérez	250.000
41489112	Microsoft	Sonia Martínez	235.000
11400200	Oracle	Ana García	280.000

Para transformar la tabla en otras que sí estén en 2FN, se crea:

1. Una tabla con la clave y todos los atributos que dependen de toda la clave. La clave sigue siendo la misma.
2. Otra tabla con la parte de la clave que tiene dependencias, junto con los atributos no primarios implicados. La clave de esta nueva tabla será esa parte de la clave.

*Ejemplo:* Transformamos la tabla en otras dos, quedando:

EMPLEADOS(DNI, Empresa, Sueldo)

PERSONAS(DNI, Nombre)

Observamos que las dos tablas se encuentran ya en 2FN, y que se mantienen las dependencias.

Es interesante hacer notar que, si una tabla está en 1FN, y su clave primaria sólo está formada por un único atributo, entonces también está en 2FN.

### 8.4.4. Tercera forma normal

Una tabla se dice que está en 3FN si y sólo si se cumplen dos condiciones:

1. Está en 2FN.
2. Todos los atributos no primarios dependen funcionalmente de la clave, y no de otros atributos no primarios.

*Ejemplo:* La tabla ALUMNOS, que contiene tres atributos que dan información sobre un alumno: NUMMAT (número de matrícula), GRUASI (grupo asignado), AULGRU (aula correspondiente al grupo). La clave es NUMMAT. Se parte de estas condiciones:

- Un alumno sólo tiene asignado un grupo.
- A un grupo siempre le corresponde un único aula.

En tal caso, tenemos que:

NUMMAT → GRUASI

NUMMAT → AULGRU

#### 8.4. Normalización de tablas relacionales

GRUASI → AULGRU

Y como vemos, un atributo no primario (AULGRU), depende funcionalmente de otro que también es atributo no primario (GRUASI). Es decir, AULGRU se puede conocer de forma única mediante la clave, y también mediante el atributo GRUASI (que no es clave ni forma parte de la clave).

NUMMAT	GRUASI	AULGRU
9941	C	P-1
8030	D	P-8
1025	C	P-1
1026	C	P-1
2077	D	P-8

Para transformar la tabla en otras que sí estén en 3FN, se crea:

1. Una tabla con la clave y todos los atributos no primarios que dependen únicamente de la clave (es decir, que no dependan también de otros atributos no primarios).
2. Otra tabla con los atributos que dependen de otro atributo no primario, incluyendo éste último (que será la clave primaria de la nueva tabla).

*Ejemplo:* La tabla ALUMNOS se descompone en otras dos, de la siguiente forma:

ALUMNOS(NUMMAT, GRUASI)

GRUPOS(GRUASI, AULGRU)

Observamos que las dos tablas se encuentran ya en 3FN, y que se mantienen las dependencias.

#### 8.4.5. Ejercicios

##### Ejercicio 1

CLIENTES(DNI, Nombre, Dirección, Teléfono, Población, Provincia, Código Postal)

*Se pide:* Construir una descomposición sin pérdidas que conserve las dependencias y que esté en 3FN.

##### Ejercicio 2

FACTURA(Número, Fecha, Cód. Cliente, Total, Cód. Artículo, Denominación, Cantidad, Precio, Importe)

*Se pide:* construir una descomposición sin pérdidas que conserve las dependencias y que esté en 3FN.

## 8. Modelos de datos

### Ejercicio 3

ESTUDIANTE(Número Estudiante, Nombre, Fecha Nacimiento, Número Curso, Nombre Curso, Nota, Número Profesor, Nombre Profesor)

Si la fila (E1, Pérez, 06/02/70, C1, Informática, Notable, P1, Rodríguez) existe en la tabla, significa que el estudiante número E1, de nombre Pérez nacido el 06/02/70, realizó el curso número C1, llamado Informática, y obtuvo una calificación de Notable. Su profesor tenía el número P1, y se llamaba Rodríguez. Además, tenemos las siguientes dependencias:

- Número Estudiante  $\rightarrow$  Nombre Estudiante | Fecha Nacimiento | Número Profesor
- Número Curso  $\rightarrow$  Nombre Curso
- Número Estudiante + Número Curso  $\rightarrow$  Nota
- Número Profesor  $\rightarrow$  Nombre Profesor

Y se sabe que un estudiante puede ir a varios cursos, y que a un curso pueden ir varios estudiantes.

*Se pide:* Dar la clave primaria de la tabla, estudiar los problemas que presenta dicha tabla, y normalizarla hasta la 3FN.

### Ejercicio 4

PROFESORES(DNI Profesor, Nombre Profesor, DNI Alumno, Nota)

Se cumple lo siguiente:

- DNI Profesor  $\rightarrow$  Nombre Profesor
- DNI Profesor + DNI Alumno  $\rightarrow$  Nota

*Se pide:* Dar la clave primaria y normalizarla hasta la 3FN.

### Ejercicio 5

NOMENCLATURA(Producto, Descripción, Compuesto, Cantidad)

Si la fila (P1, Coche, P3, 4) existe en la tabla, significa que para la fabricación de una unidad de producto P1, cuya descripción es Coche, se necesita el producto P3 en la cantidad 4. Además, tenemos las siguientes dependencias:

- Producto  $\rightarrow$  Descripción
- Producto + Compuesto  $\rightarrow$  Cantidad



#### 8.4. Normalización de tablas relacionales

Y se sabe que un producto puede estar formado por varios compuestos, y que un mismo compuesto puede usarse para fabricar varios productos.

*Se pide:* Dar la clave primaria de la tabla, estudiar los problemas que presenta dicha tabla, y normalizarla hasta la 3FN.

