

# Anomalías y formas normales

Álvaro González Sotillo

8 de noviembre de 2018

## Índice

|                             |   |
|-----------------------------|---|
| 1. Introducción             | 1 |
| 2. Anomalías                | 1 |
| 3. Dependencias funcionales | 3 |
| 4. Normalización            | 5 |
| 5. Referencias              | 8 |

## 1. Introducción

- El diseño de base de datos es necesario para evitar anomalías
  - De inserción
  - De modificación
  - De borrado
- Las anomalías hacen el uso de una base de datos
  - Engorroso
  - Lento
  - Propenso a errores
- A medio plazo, las anomalías hacen *fracasar* una base de datos

## 2. Anomalías

### 2.1. Cambios de la BD (diseño E-R)

- Desde el punto de vista del diseño E-R, un **cambio** en la base de datos es:
  - Modificación de atributos de una instancia de entidad
  - Modificación de atributos de una instancia de interrelación
  - Creación/borrado de una nueva instancia de entidad
  - Creación/borrado de una nueva instancia de interrelación

---

## 2.2. Cambios de la BD (relaciones)

- Desde el punto de vista del diseño relacional, un **cambio** en la base de datos es:
  - Modificación de atributos de una fila
  - Creación de una fila
  - Borrado de una fila

## 2.3. Anomalía

- Decimos que hay una anomalía cuando **un** cambio desde el punto de vista E-R implica **más de un** cambio en la base de datos relacional
  - Anomalía de inserción
  - Anomalía de borrado
  - Anomalía de modificación

### 2.3.1. Anomalía de inserción/borrado

| Idempleado | NombreP        | Puesto         | Salario | Centro             | DirecciónC  | Teléfono |
|------------|----------------|----------------|---------|--------------------|-------------|----------|
| 123A       | Ana Almansa    | Profesor       | 20.000  | Informática        | Complutense | 123      |
| 456B       | Bernardo Botín | Administrativo | 15.000  | Matemáticas        | Complutense | 456      |
| 789C       | Carlos Crespo  | Catedrático    | 30.000  | CC . Empresariales | Somosaguas  | 789      |
| 012D       | David Díaz     | Ayudante       | 10.000  | Informática        | Complutense | 123      |

- ¿Cómo inserto un empleado que aún no tiene centro de trabajo?
- ¿Cómo inserto un centro de trabajo que aún no tiene empleados?

### 2.3.2. Anomalía de modificación

| Idempleado | NombreP        | Puesto         | Salario | Centro             | DirecciónC  | Teléfono |
|------------|----------------|----------------|---------|--------------------|-------------|----------|
| 123A       | Ana Almansa    | Profesor       | 20.000  | Informática        | Complutense | 123      |
| 456B       | Bernardo Botín | Administrativo | 15.000  | Matemáticas        | Complutense | 456      |
| 789C       | Carlos Crespo  | Catedrático    | 30.000  | CC . Empresariales | Somosaguas  | 789      |
| 012D       | David Díaz     | Ayudante       | 10.000  | Informática        | Complutense | 123      |

- ¿Cómo muevo la facultad de informática a Somosaguas?

### 2.3.3. Origen de las anomalías

- Redundancia de las interrelaciones
- Mala elección de la clave
- Mala elección de la cardinalidad de las interrelaciones

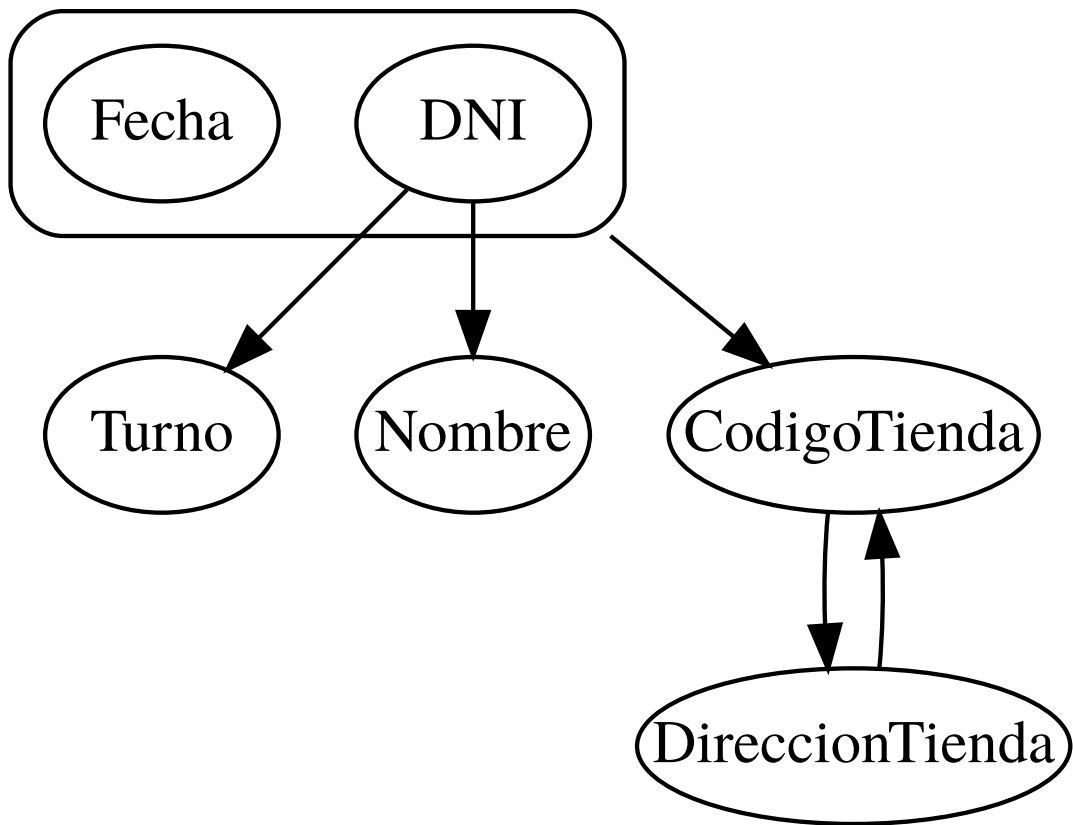
---

### 3. Dependencias funcionales

- Decimos que un campo B depende funcionalmente de otro A si, al conocer el valor de A, conocemos el valor de B.
- Las dependencias pueden ser
  - Solo de un campo a otro campo:  $A \rightarrow B$
  - De varios campos a un campo:  $(A, B) \rightarrow C$
- Las dependencias se extraen del **significado** de los datos
  - Sólo el cliente/usuario final conoce dicho significado
  - El diseñador de bases de datos debe extraer dicho conocimiento y formalizarlo

#### 3.1. Ejemplo

- Se desea una base de datos que guarde los atributos (DNI, Nombre, CódigoTienda, DirecciónTienda, Fecha, Turno)
- Los empleados tienen un turno, y cada día pueden trabajar en una tienda.
- No hay dos tiendas en la misma dirección
- Dependencias:
  - $DNI \rightarrow Nombre$
  - $DNI \rightarrow Turno$
  - $CódigoTienda \rightarrow DirecciónTienda$
  - $DirecciónTienda \rightarrow CódigoTienda$
  - $(DNI, Fecha) \rightarrow CódigoTienda$
- Hay más dependencias, pero pueden deducirse de las anteriores
  - Por ejemplo  $(DNI, Fecha, Turno) \rightarrow DirecciónTienda$

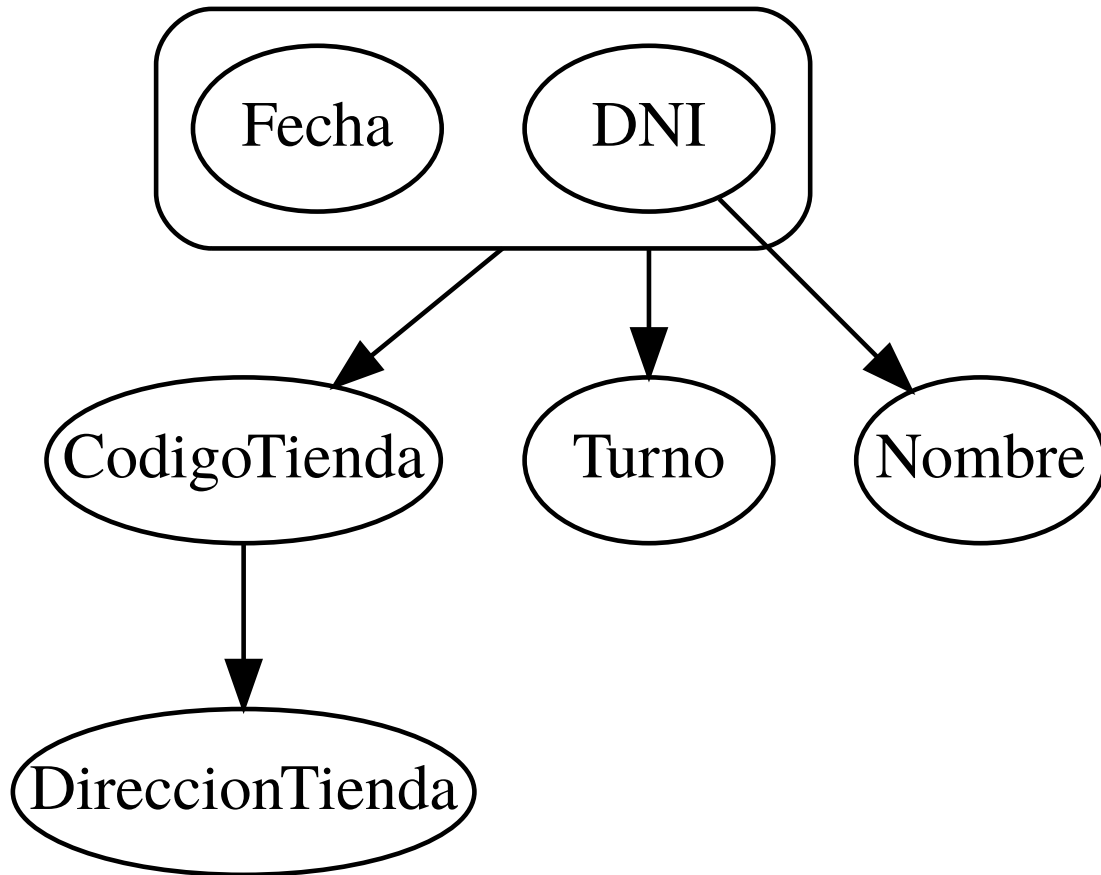


### 3.2. Ejercicio

- Se desea una base de datos que guarde los atributos (DNI, Nombre, CódigoTienda, DirecciónTienda, Fecha, Turno)
- Los empleados no tienen un turno fijo, y cada día pueden trabajar en una tienda en un turno
- En una misma dirección puede haber varias tiendas

---

### 3.2.1. Solución



## 4. Normalización

- Procedimiento para eliminar anomalías en la base de datos
- Es formal: a partir de las *dependencias funcionales* es un procedimiento automático

### 4.1. Clave de una relación

- La clave de una relación es el conjunto mínimo de campos que es imposible que se repitan en la relación
  - Coinciden con los campos de los que depende funcionalmente el resto de campos
- Sirven para *localizar* cada fila de la relación
- Si hay varias claves posibles
  - Todas ellas son **claves candidatas**
  - Se elige una como **clave primaria**

---

## 4.2. Primera forma normal (1FN)

- Una relación está en **1FN** si no tiene atributos multivaluados

### 4.2.1. Ejemplo 1FN

| IDCliente | Nombre | Apellido | Teléfono                   |
|-----------|--------|----------|----------------------------|
| 123       | Rachel | Ingram   | 555-861-2025               |
| 456       | James  | Wright   | 555-403-1659, 555-776-4100 |
| 789       | Cesar  | Dure     | 555-808-9633               |

- La clave es IDCLIENTE
- La relación se parte en dos:
  - Una con todos los atributos que respetan 1FN
  - Otra con la clave IDCLIENTE y el atributo que no respeta 1FN

| IDCliente | Nombre | Apellido |
|-----------|--------|----------|
| 123       | Rachel | Ingram   |
| 456       | James  | Wright   |
| 789       | Cesar  | Dure     |

| IDCliente | Teléfono     |
|-----------|--------------|
| 123       | 555-861-2025 |
| 456       | 555-403-1659 |
| 456       | 555-776-4100 |
| 789       | 555-808-9633 |

## 4.3. Segunda forma normal (2FN)

- Una relación está en **2FN** si
  - Está en 1FN
  - Todos los atributos que no son parte de una clave candidata dependen de la totalidad de las claves candidatas
- Como consecuencia, si está en 1FN y la clave candidata es de un atributo, está en 2FN

### 4.3.1. Ejemplo 2FN

| Empleado | Habilidad       | Lugardetrabajo    |
|----------|-----------------|-------------------|
| Jones    | Mecanografía    | 114 Main Street   |
| Jones    | Taquigrafía     | 114 Main Street   |
| Jones    | Tallado         | 114 Main Street   |
| Bravo    | Limpieza ligera | 73 Industrial Way |
| Ellis    | Alquimia        | 73 Industrial Way |
| Ellis    | Malabarismo     | 73 Industrial Way |
| Harrison | Limpieza ligera | 73 Industrial Way |

- La clave es (Empleado, Habilidad)

- 
- Lugar de trabajo depende de Empleado
  - Hay anomalías de modificación
  - Hay redundancias de datos
  - Para conseguir 2FN se parte la relación en dos
    - Se deja la clave y todos los atributos que dependen totalmente de ella
    - Se crea una relación con el atributo que depende de una parte de la clave, con esa parte de la clave

| Empleado | Lugar actual de trabajo |
|----------|-------------------------|
| Jones    | 114 Main Street         |
| Bravo    | 73 Industrial Way       |
| Ellis    | 73 Industrial Way       |
| Harrison | 73 Industrial Way       |

| Empleado | Habilidad       |
|----------|-----------------|
| Jones    | Mecanografía    |
| Jones    | Taquigrafía     |
| Jones    | Tallado         |
| Bravo    | Limpieza ligera |
| Ellis    | Alquimia        |
| Ellis    | Malabarismo     |
| Harrison | Limpieza ligera |

#### 4.4. Tercera forma normal (3FN)

- Una relación está en **3FN** si
  - Está en 2FN
  - No hay dependencias transitivas de la clave

##### 4.4.1. Ejemplo 3FN

| Torneo               | Año  | Ganador        | Fecha de nacimiento del ganador |
|----------------------|------|----------------|---------------------------------|
| 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             |

- La clave es (Torneo, Año)
- Fecha de nacimiento depende de Ganador, que depende de la clave
- Hay anomalías de modificación
- Para conseguir 3FN se parte la relación en 2
  - Se deja la clave y los atributos sin dependencia transitiva
  - Se crea una relación que tiene como clave y campos la dependencia transitiva

---

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

| Ganador        | Fechadenacimiento        |
|----------------|--------------------------|
| Chip Masterson | 14 de marzo de 1977      |
| Al Fredrickson | 21 de julio de 1975      |
| Bob Albertson  | 28 de septiembre de 1968 |

## 5. Referencias

- Formatos:
  - [Transparencias](#)
  - [PDF](#)
- Creado con:
  - [org-reveal](#)
  - [Latex](#)