

# Tema 3

Modelos de datos.  
El modelo de datos relacional.

# Contenidos

- Definición de modelo de datos
- El modelo de datos relacional
  - Estructura de datos
  - Restricciones de integridad
- Otros modelos de datos

# Contenidos

- Definición de modelo de datos
- El modelo de datos relacional
  - Estructura de datos
  - Restricciones de integridad
- Otros modelos de datos

# Definición de modelo de datos

- Proceso de transformación
- a) Mundo real
  - Delimitación objetivos
  - Selección de datos
  - Hipótesis semánticas
  - Organización de los datos a almacenar
- b) Esquema Inicial
  - Datos operativos
    - Atributos
    - Conexiones
    - Restricciones

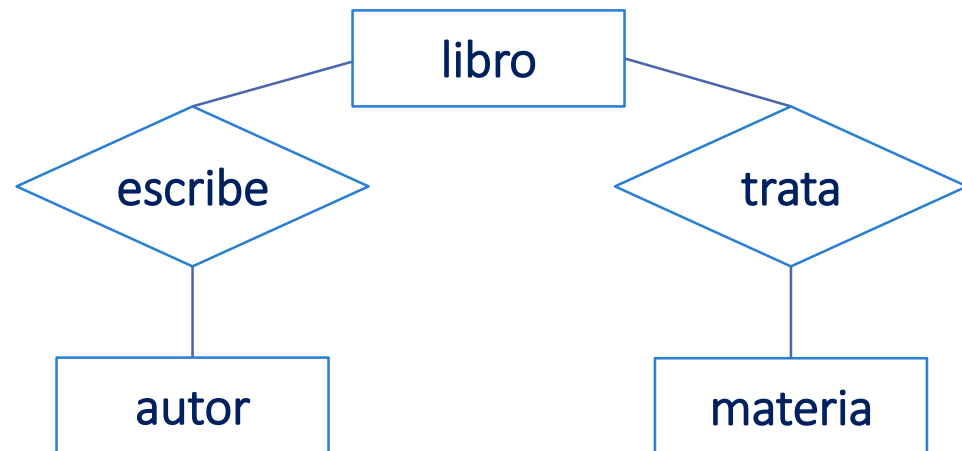
Problema tipo

a) Biblioteca



b) Resultado

- Libro: título, isbn, editorial....
- Autor: nombre, nacionalidad,....
- Materia: código, descripción....



# Definición de modelo de datos

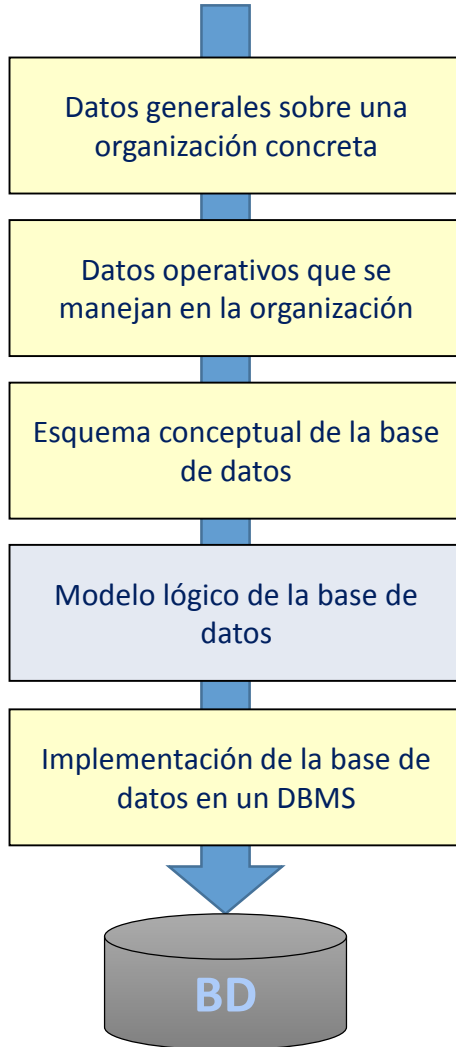
- Definición formal

- Mecanismo formal para representar y manipular información de manera general y sistemática.
- Debe constar de:
  - 1.- Notación para describir datos
  - 2.- Notación para describir operaciones
  - 3.- Notación para describir reglas de integridad

- Historia

- 1<sup>er</sup> Modelo es el relacional (Codd).
- Se recuperan los modelos basados en grafos (1974).
- El modelo E/R (Chen, 1975), otros modelos semánticos.
- Modelos orientados a objetos (1983, 1986,...).
- Modelos lógicos (1986...).

# Definición de modelo de datos



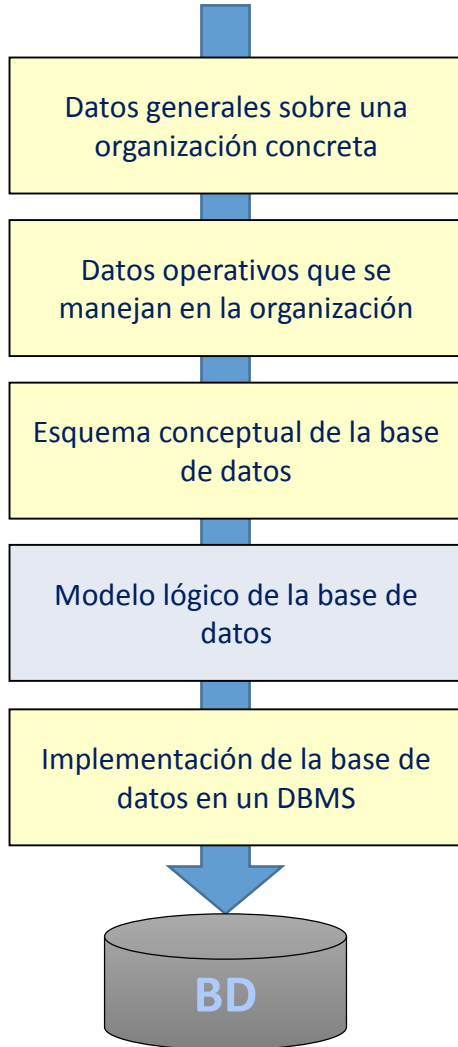
## Modelado lógico:

Trasladamos a un esquema lógico en función de una estructura implementable.

Título	ISBN	Editorial	...
Introducción a las BD	1234-1234	Thomson	...
Cálculo para todos	4321-4321	Delta	...
...	...	...	...

## • Tabla LIBROS

# Definición de modelo de datos



- Lo implementamos en un sistema comercial:

```
CREATE TABLE LIBROS (  
    titulo char(45) NOT NULL,  
    ISBN char(10) PRIMARY KEY,  
    editorial char(30) REFERENCES ...  
);
```

# Definición de modelo de datos

- Necesidad de modelos de datos:
  - Cada esquema se describe utilizando un lenguaje de definición de datos.
  - Este lenguaje es de muy bajo nivel, está muy ligado al SGBD.
  - Hacen falta otros mecanismos de más alto nivel que permitan describir datos de una forma no ambigua y entendible por los usuarios en cada paso del proceso de implantación.



# Definición de modelo de datos

- Objetivo:
  - Describir modelos que representen los datos y los describan de una forma entendible y manipulable.
  - En relación con la Arquitectura ANSI/SPARC:
    - Nivel Externo:
      - Modelo de datos externo
    - Nivel Conceptual:
      - Modelo de datos conceptual
    - Nivel Interno:
      - Modelo de datos interno

# Definición de modelo de datos

- Clasificación:
  - Basados en registros
  - Basados en objetos
  - Físicos
- Utilización:
  - Los dos primeros:
    - Nivel externo y conceptual
  - Físicos:
    - Nivel interno

# Definición de modelo de datos

- Modelos de datos basados en registros:
  - Modelo de datos jerárquico
  - Modelo de datos en red
  - Modelo de datos relacional (Codd, 1969)

# Contenidos

- Definición de modelo de datos
- El modelo de datos relacional
  - Estructura de datos
  - Restricciones de integridad
- Otros modelos de datos

# El modelo de datos relacional

## Estructura de datos

- El modelo de datos relacional organiza y representa los datos en forma de tablas o relaciones.
- Base de datos relacional: colección de tablas cada una de las cuales tiene un nombre único.

# El modelo de datos relacional

## Estructura de datos

The diagram illustrates a relational database table with the following structure and data:

Id_trabajador	Nombre	Tarifa_hr	Tipo_de_oficio	Id_supv
1235	M. López	12,50	Electricista	1311
1412	J.L. Calvo	13,75	Fontanero	1520
2920	N. Marín	10,00	Carpintero	Nulo
3231	O. Pons	17,40	Albañil	Nulo
1540	M.A. Vila	11,75	Fontanero	Nulo
1311	J.C. Cubero	15,50	Electricista	Nulo
3001	D. Sánchez	8,20	Albañil	3231

Annotations in the diagram:

- Atributo**: Points to the header row (attributes).
- Dominio**: Points to the first column (Id\_trabajador).
- Tupla**: Points to a single row of data.

Figura 1: Tabla Trabajadores

# El modelo de datos relacional

## Estructura de datos

- Algunos conceptos:
  - Esquema de una base de datos relacional:
    - Colección de esquemas de relaciones junto con restricciones de integridad.
  - Instancia o estado de una base de datos:
    - Colección de instancias de relaciones que verifican las restricciones de integridad.
  - Base de datos relacional:
    - Instancia de una base de datos junto con su esquema.

# El modelo de datos relacional

## Estructura de datos

- Introducido por E.F. Codd en 1970.
- El modelo relacional abarca tres ámbitos distintos de los datos:
  - Las estructuras para almacenarlos: El usuario percibe la información de la base de datos estructurada en tablas.
  - La integridad: Las tablas deben satisfacer ciertas condiciones que preservan la integridad y la coherencia de la información que contienen.
  - Consulta y manipulación: Los operadores empleados por el modelo se aplican sobre tablas y devuelven tablas.
- La tabla es la estructura lógica de un sistema relacional. A nivel físico, el sistema es libre de almacenar los datos en el formato más adecuado (archivo secuencial, archivo indexado, listas con apuntadores,...).



# El modelo de datos relacional

## Estructura de datos

- Atributo: Cualquier elemento de información susceptible de tomar valores. Notación:  $A_i$ ,  $i=1,2,\dots$
- Dominio: Rango de valores donde toma sus datos un atributo. Se considera finito. Notación:  $D_i$ ,  $i=1,2,\dots$
- Relación: Dados los atributos  $A_i$ ,  $i=1,2,\dots,n$  con dominios  $D_i$ ,  $i=1,2,\dots,n$ , no necesariamente distintos, definimos relación asociada a  $A_1..A_n$ , y lo notaremos por  $R(A_1..A_n)$ , a cualquier subconjunto del producto cartesiano  $D_1 \times D_2 \times \dots \times D_n$ .

# El modelo de datos relacional

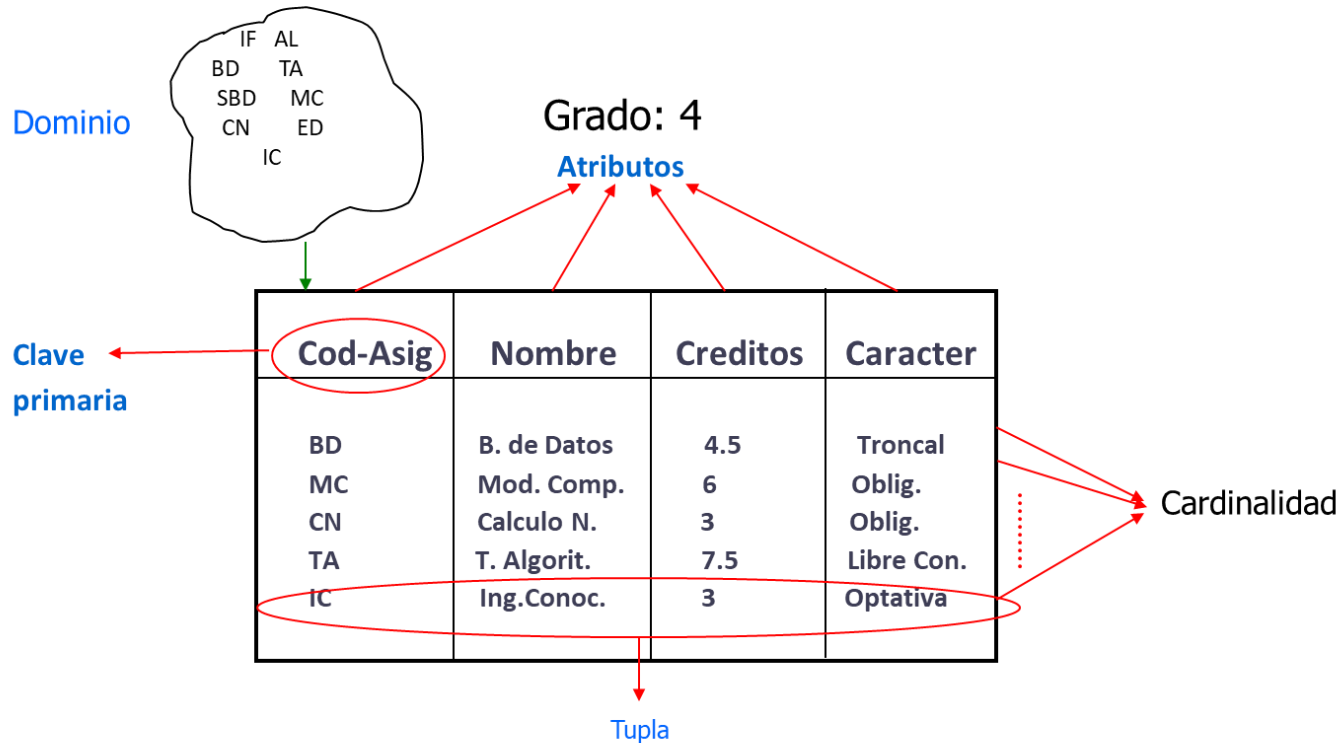
## Estructura de datos

- Tupla
  - Cada una de las filas de una relación.
- Cardinalidad de una relación
  - Número de tuplas que contiene.
  - Es variable en el tiempo.
- Esquema de una relación R
  - Atributos  $A_1:D_1, \dots, A_n:D_n$
- Grado de una relación
  - Número de atributos de su esquema.
  - Invariable en el tiempo
- Instancia de una relación
  - Conjunto de tuplas  $\{(x_1, x_2, \dots, x_n)\} \subseteq D_1 \times D_2 \times \dots \times D_n$  que la componen en cada momento

# El modelo de datos relacional

## Estructura de datos

### Ejemplo: Asignaturas de una titulación



# El modelo de datos relacional

## Estructura de datos

- Propiedades

- Condición de normalización

- Todos los valores de los atributos de una relación son atómicos
      - Valor atómico es un valor no estructurado
    - Cuando una relación cumple la primera condición de normalización se dice que está en Primera Forma Normal.

- Consecuencias

- No hay valores tipo conjunto
    - No hay valores tipo registro
    - No hay valores tipo tablas

- Problema

- Todas las representaciones son extensivas: no se puede representar directamente información del tipo “el valor del atributo asignaturas de un alumno es : (FBD, ALG, LD)”

# El modelo de datos relacional

## Estructura de datos

- Consecuencias de la definición
  - No hay tuplas duplicadas
    - Por la definición conjuntista de relación
  - No hay orden en las filas ni en los atributos
    - Al no estar ordenados ni los atributos ni las filas (conjuntos) el acceso es por Nombre de Atributo y Valor
  - Varias instancias representan la misma relación

A	B	C	D	E
<i>a1</i>	<i>b1</i>	<i>c1</i>	<i>d1</i>	<i>e1</i>
<i>a1</i>	<i>b2</i>	<i>c2</i>	<i>d2</i>	<i>e1</i>
<i>a2</i>	<i>b1</i>	<i>c3</i>	<i>d3</i>	<i>e1</i>
<i>a2</i>	<i>b1</i>	<i>c4</i>	<i>d3</i>	<i>e1</i>
<i>a3</i>	<i>b2</i>	<i>c5</i>	<i>d1</i>	<i>e1</i>

A	B	C	D	E
<i>a2</i>	<i>b1</i>	<i>c4</i>	<i>d3</i>	<i>e1</i>
<i>a2</i>	<i>b1</i>	<i>c3</i>	<i>d3</i>	<i>e1</i>
<i>a1</i>	<i>b2</i>	<i>c2</i>	<i>d2</i>	<i>e1</i>
<i>a3</i>	<i>b2</i>	<i>c5</i>	<i>d1</i>	<i>e1</i>
<i>a1</i>	<i>b1</i>	<i>c1</i>	<i>d1</i>	<i>e1</i>

# El modelo de datos relacional

## Estructura de datos

- Representación Física    Representación Intuitiva    Modelo Matemático
  - Archivo secuencial    ↔    Tabla    ↔    Relación
  - Registros    ↔    Filas    ↔    Tuplas
  - Campos    ↔    Columnas    ↔    Valores atributos
- Esquema de una base de datos relacional
  - Colección de esquemas de relaciones junto con restricciones de integridad
- Instancia o estado de una base de datos
  - Colección de instancias de relaciones que verifican las restricciones de integridad
- Base de datos relacional
  - Instancia de una base de datos junto con su esquema

# El modelo de datos relacional

## Estructura de datos

- Notación a utilizar

- Relación:  $R, S, T, \dots$
- Atributos:  $A, B, \dots$
- Esquema de relación:  $R[A_1, A_2, \dots, A_n]$
- Instancia de relación  $R$ :  $r, \dots$
- Tuplas de una instancia:  $x_1, x_2, \dots \in r$
- Valor de un atributo  $A_i$  en una tupla  $x_j$ :  $x_j[A_i]$  ó  $A_{ij}$

# El modelo de datos relacional

## Estructura de datos

- Algunas veces no se conoce el valor de un atributo para una determinada tupla. En esos casos a ese atributo de esa tupla se le asigna un valor nulo.
  - Un valor nulo puede ser un valor desconocido.
  - Un valor nulo puede ser un atributo no aplicable.
  - En cualquier caso, ese valor es un valor más de todos los dominios de la base de datos.



# El modelo de datos relacional

## Restricciones de integridad

- Restricciones o reglas de integridad son condiciones para preservar la semántica de una base de datos
- Específicas del problema:
  - $0 \leq \text{edad} \leq 100$
  - $\text{créditos} > 0$
  - $\text{carácter} \in (\text{'troncal'}, \text{'obligatoria'}, \text{'optativa'}, \dots)$
- Propias del papel de los atributos en el esquema:
  - $\text{imparte.NRP} \in \text{profesor.NRP}$   
(un profesor inexistente no puede impartir una asignatura)
  - $\text{cod\_asig} \neq \text{nulo}$   
(siempre debe conocerse el código de una asignatura)

# El modelo de datos relacional

## Restricciones de integridad

- Superclaves y claves (candidatas y primarias)
  - Superclave: Cualquier conjunto de atributos que identifica unívocamente a cada tupla de una relación.
- Clave (Candidata): superclave minimal.
  - Ejemplo:
    - Relación Trabajadores
      - {id\_trabajador, nombre} – superclave
      - No es minimal: no es clave de la relación
    - Por ejemplo, en la relación Asignaturas, el conjunto de atributos {Cod\_Asig, Nombre} identifica unívocamente cada tupla. Sin embargo, no es minimal y no puede considerarse como una clave. Cod\_Asig por sí solo, es una clave.
- De entre las candidatas (si hubiera más de una), hay que elegir una como principal que se denomina Clave Primaria.
  - Criterio de selección: Tamaño, significado, capacidad para recordarla, fusión con otras tablas,...

# El modelo de datos relacional

## Restricciones de integridad

- Clave candidata y primaria (definición formal).
  - Sea  $R[A_1, A_2, \dots, A_n]$ ,  $PK \subseteq \{A_1, A_2, \dots, A_n\}$  se denomina clave candidata sii:
    - Unicidad:  $\forall r$  instancia de  $R$  y  $\forall t_1, t_2 \in r$   $t_1 \neq t_2 \Rightarrow t_1[PK] \neq t_2[PK]$
    - Minimalidad: No existe  $PK' \subset PK$  que verifique la unicidad.
  - Es decir, una clave candidata es un atributo o conjunto de atributos que identifica a cada tupla en la relación y que, además, no existe un subconjunto de ellos que también identifique a cada tupla de la relación.
  - Una clave primaria es una clave candidata elegida por el diseñador.
  - Si  $PK$  verifica la unicidad y no la minimalidad se denomina superclave.

# El modelo de datos relacional

## Restricciones de integridad

- Conceptos generales
  - Condiciones de integridad:
    - Normas que mantienen la corrección semántica de una base de datos.
  - Nos centramos en Integridad Genérica: depende del papel que juegue un atributo en el diseño de la tabla.
    - Son metarreglas (generan reglas de integridad aplicadas a una base de datos concreta).
    - Existen la integridad de entidad y la integridad referencial.

# El modelo de datos relacional

## Restricciones de integridad

- Integridad de entidad
  - No se debe permitir que una entidad sea representada en la base de datos si no se tiene una información completa de los atributos que son clave primaria de la entidad:
    - La clave primaria, o una parte de la misma, no puede ser un valor nulo.

*Un atributo que forma parte de la clave primaria de una tupla en una relación no puede tener un valor nulo.*

# El modelo de datos relacional

## Restricciones de integridad

- Clave externa (ajena): conjunto de atributos en una relación que es una clave en otra (o incluso en la misma) relación.
- Podemos ver una clave externa como un conjunto de atributos de una relación cuyos valores en las tuplas deben coincidir con valores de la clave primaria de las tuplas de otra relación.

# El modelo de datos relacional

## Restricciones de integridad

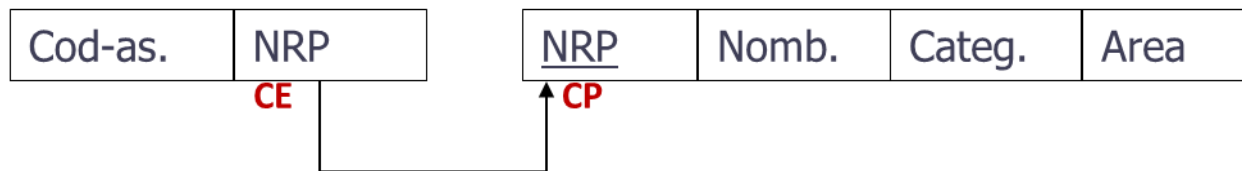
- Formalmente:

- Clave externa

- Sean  $R[A_1, A_2, \dots, A_n]$ , y  $PK \subseteq \{A_1, A_2, \dots, A_n\}$  su clave primaria, sea  $S[B_1, B_2, \dots, B_n]$ , y  $FK \subseteq \{B_1, B_2, \dots, B_n\}$  de manera que  $\text{card}(PK) = \text{card}(FK)$ . Fk es clave externa de S con respecto a R si se verifica:
      - $\forall r$  instancia de R y  $\forall s$  instancia de S,  $\forall x \in s \Rightarrow \exists y \in r / x[FK] = y[PK]$
      - Es decir, el “dominio activo” de FK debe estar incluido en el “dominio activo” de PK para cualquier instancia de la base de datos.

Imparte

Profesor



# El modelo de datos relacional

## Restricciones de integridad

- Integridad referencial:
  - Una base de datos en la que todos los valores no nulos de una clave externa referencian valores reales de la clave referenciada en la otra relación cumple la regla de integridad referencial.

*Si una relación incluye una clave externa conectada a una clave primaria, el valor de la clave externa debe ser, bien igual a un valor ya existente en el dominio activo de la clave primaria, o bien completamente nulo (si la semántica lo permite).*

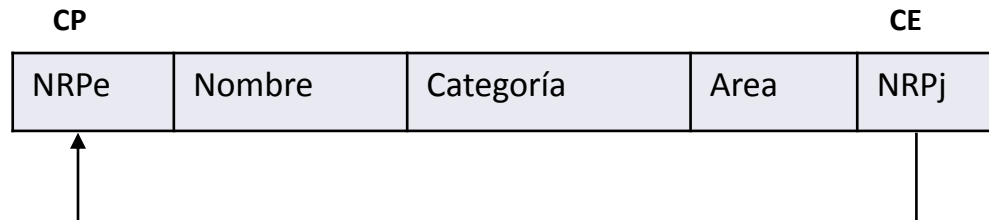
- La integridad referencial mantiene las conexiones en las bases de datos relacionales.



# El modelo de datos relacional

## Restricciones de integridad

- La integridad referencial
  - Puede haber más de una clave externa en una relación
  - Puede haber una clave externa a la clave primaria de la propia relación
- Ejemplo



# El modelo de datos relacional

## Restricciones de integridad

- EL SGBD debe encargarse de mantener las siguientes restricciones:
  - La unicidad de la clave primaria y de las claves candidatas:
    - Frente a operaciones de Inserción y Actualización, el SGBD debe rechazar los valores introducidos que sean iguales a los presentes en la BD para los atributos que el diseñador ha definido como clave primaria y como claves candidatas.

# El modelo de datos relacional

## Restricciones de integridad

- La restricción de integridad de identidad:
  - Frente a operaciones de Inserción y Actualización, el SGBD debe rechazar las modificaciones que vulneren la unicidad en la clave primaria y/o que asignen un valor NULO a algún atributo de la clave primaria.

# El modelo de datos relacional

## Restricciones de integridad

- Integridad referencial

- En inserción:

- Rechazar la tupla insertada si el valor de la clave externa no concuerda en la relación referenciada para alguna tupla en el valor su clave primaria.
    - Si el valor para la clave externa es NULO y el diseño no lo permite habrá de rechazar también esa inserción.

- En actualización:

- Si se actualiza la clave externa: rechazar la modificación si se produce alguna de las circunstancias descritas en punto anterior.
    - Si se actualiza la clave primaria de la relación referenciada: Actualizar en cadena las claves externa que la referencien (o impedir la actualización mientras existan referencias a valor anterior).

- En borrado:

- Si se borra la clave primaria en la relación referenciada: borrado en cadena de todas las tuplas que la referencian o poner valor nulo en la clave externa de todas esas tuplas.

# Contenidos

- Definición de modelo de Datos
- El modelo de datos relacional
  - Estructura de datos
  - Restricciones de integridad
- Otros modelos de datos