
1.- Presentación de bases de datos.

Ideas generales y ejemplos (por ejemplo de alumnos, materias, profesores, etc.)

2.- Información en sistemas de archivos tradicionales frente a Sistemas de Gestión de Bases de Datos (SGBD)

3.- Nomenclatura básica (conjunto, tupla, etc.) y estructuras de datos usadas en Bases de Datos.

4.- Modelos de Bases de Datos.

El modelo relacional.

5.- Desarrollo de un ejemplo concreto de aplicación.

En la vida diaria se hace un uso habitual de información, mucha de la cuál está gestionado por Sistemas de Gestión de Bases de Datos:

Compras pagadas con tarjeta

Reservas de billetes de avión

Reservas por internet de habitaciones de hotel

En la universidad: temas de alumnos, trabajadores, cursos a través del campus virtual, etc.

Préstamo de libros en bibliotecas.

Seguros de coche, hogar, etc.

Fabricación: producción, inventario, pedidos, proveedores, etc.

La información contenida puede ser muy grande. En la actualidad mucha información de nuestra vida diaria está contenida y gestionada por bases de datos (Twitter, Facebook, etc.)

Ejemplo sistema Información para una universidad

Ejemplos de programas de aplicación sobre la información contenida:

Añadir nuevos estudiantes, profesores, titulaciones o asignaturas ofertadas.

Registrar a los estudiantes en asignaturas y generar listados de matriculados por asignaturas.

Introducir notas a los estudiantes, calcular estadísticas y generar actas.

...

En las primeras etapas de la informática, gran parte de estas aplicaciones se realizaban directamente sobre sistemas de ficheros.

Un sistema de información es un conjunto de elementos de información organizados de acuerdo a ciertas normas que sirven para proporcionar información para los objetivos fijados previamente.

Las funciones básicas de todo sistema de información son:

- Recogida de datos
- Procesamiento de los datos
- Almacenamiento de los datos
- Elaboración y presentación de datos y resultados

En todo sistema de Información aparecen los siguientes elementos:

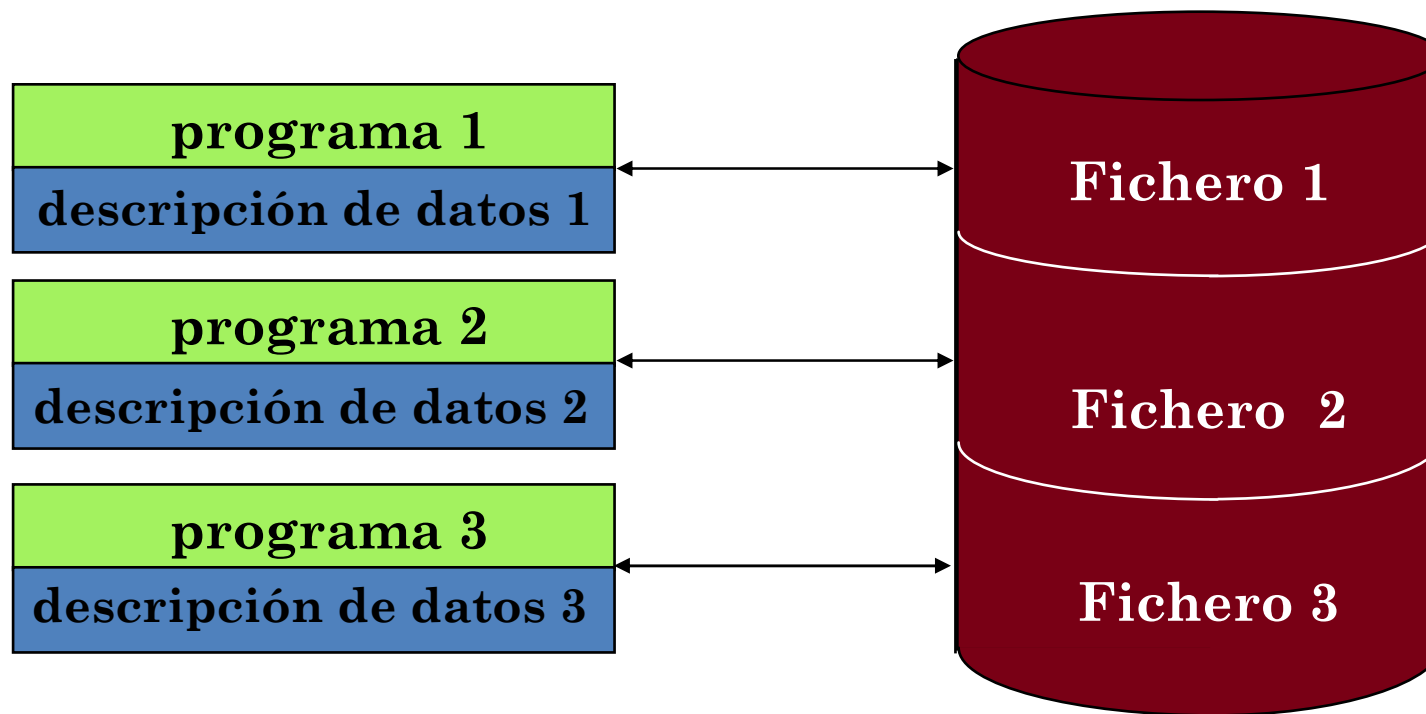
- El contenido: los **datos**, su descripción y los programas que los manipulan.
- El **equipo físico**: el sistema informático que proporciona soporte al sistema de información
- El **equipo lógico**: sistema de comunicaciones, sistema operativo, etc. En este punto aparece también el sistema de gestión de base de datos.
- El **administrador**: persona o equipo de personas responsables de garantizar la calidad y disponibilidad de los datos.
- Los **usuarios** de la información

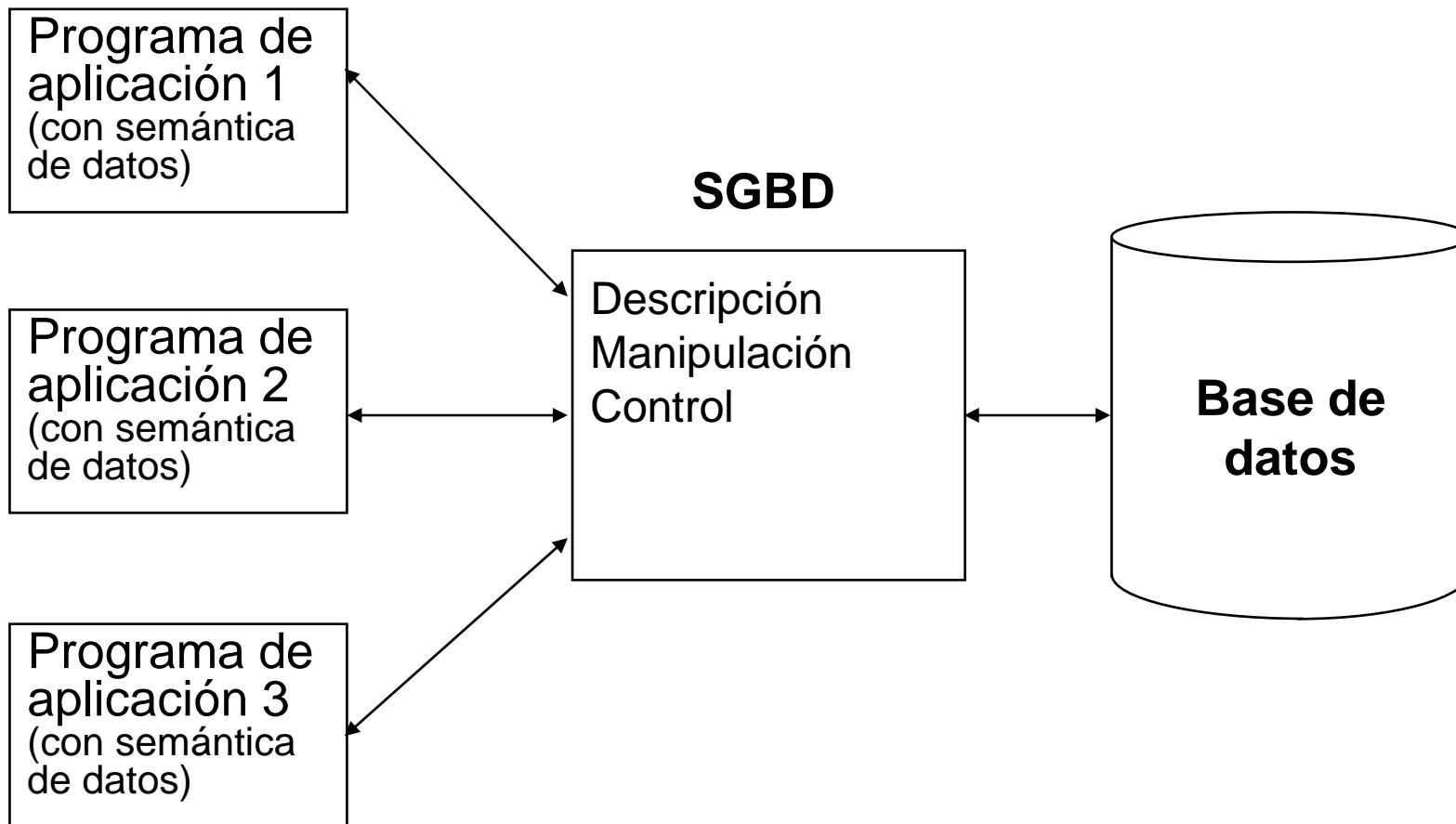
Información con sistemas de archivos

En general no se puede decir si, para una aplicación concreta, es mejor disponer de un sistema de gestión de bases de datos o utilizar la información en un sistema de archivos. Dependiendo de la información, el tratamiento presente y futuro de la misma habrá que determinar cuál es la opción más adecuada.

Vamos a plantear algunas ideas de lo que supone manejar sistemas de ficheros.

Información con sistemas de archivos





Conceptos respecto al uso de sistemas de archivos

Tipos de archivos:

ASCII o de texto

Binarios. No son legibles directamente por el usuario.

Formas de acceso:

Secuencial. Se leen y escriben los datos del mismo modo que se hace en una antigua cinta de casete. Si se quiere acceder a un dato que está hacia la mitad de un fichero, habrá que pasar primero por todos los datos anteriores.

Aleatorio. Permiten acceder directamente a un dato sin tener que pasar por todos los demás, y pueden acceder a la información en cualquier orden. Tienen la limitación de que los datos están almacenados en unas unidades o bloques que se llaman registros, y que todos los registros que se almacenan en un fichero deben ser del mismo tamaño.

Indexado. Necesitan de estructuras de datos adicionales

Problemas con los sistemas de ficheros

En general no se puede decir si, para una aplicación concreta, es mejor disponer de una sistema de gestión de bases de datos o utilizar la información en un sistema de archivos.

Vamos a plantear algunas ideas de lo que supone manejar sistemas de ficheros.

- Separación y aislamiento de los datos
- Duplicación de datos
 - Incoherencia
- Dependencia de los datos
 - Si no cambio el programa, no puedo cambiar los datos
- Formatos incompatibles
 - Distintos formatos para el almacenamiento
- Consultas fijas y proliferación de aplicaciones
- Ausencia de cierta funcionalidad:
 - Seguridad, recuperación, concurrencia

Sistemas de archivos frente a SGBD (1)

Tema	Información en ficheros	S.G.B.D.
Redundancia de datos	Puede existir una redundancia elevada de los datos	Si la base de datos está bien diseñada existirá una redundancia baja
Independencia de los datos	Cambios en el formato de los registros o en el método de acceso implica que la aplicación debe ser cambiada	Cambios semejantes no tienen ningún impacto o lo tienen muy leve
Mantenimiento de las aplicaciones	El bajo nivel de la independencia de datos implica un coste elevado de mantenimiento (cambios en los datos suponen cambios en el programa de aplicación).	La elevada independencia de los datos reduce los costes de mantenimiento debidos a la reorganización de la base de datos.

Problema: Si un dato se modifica en un fichero puede no modificarse en otro u otros.

Sistemas de archivos frente a SGBD (2)

Tema	Información en ficheros	S.G.B.D.
Controles de integridad	Si existen los debe proporcionar el programador.	Los proporciona el SGBD.
Relaciones entre datos	Los deben gestionar las aplicaciones	Es una parte integral del SGBD
Recursos consumidos	Poca CPU y poco espacio de almacenamiento adicional	Se precisa más tiempo de proceso y mayor espacio de almacenamiento
Diseño de la base de datos	Los ficheros se diseñan cuando se precisa	Se debe consumir más tiempo y esfuerzo para conseguir un buen diseño de la base de datos

Ventajas de los SGBD

- Definición y descripción del conjunto de datos única e integrada
- Control de la redundancia
- Control de accesos concurrentes
- Coherencia de los datos
- Restricción de accesos no autorizados
- Capacidad de inferencia
- Múltiples interfaces de usuario
- Garantizan las restricciones de integridad
- Copias de seguridad y recuperación frente a fallos
- Independencia de los datos respecto a los tratamientos y viceversa
- Gestión de transacciones

Una base de datos es una colección de datos:

Usada para representar información de interés para diferentes elementos de una organización.

Compartida entre diferentes aplicaciones y diferentes usuarios.

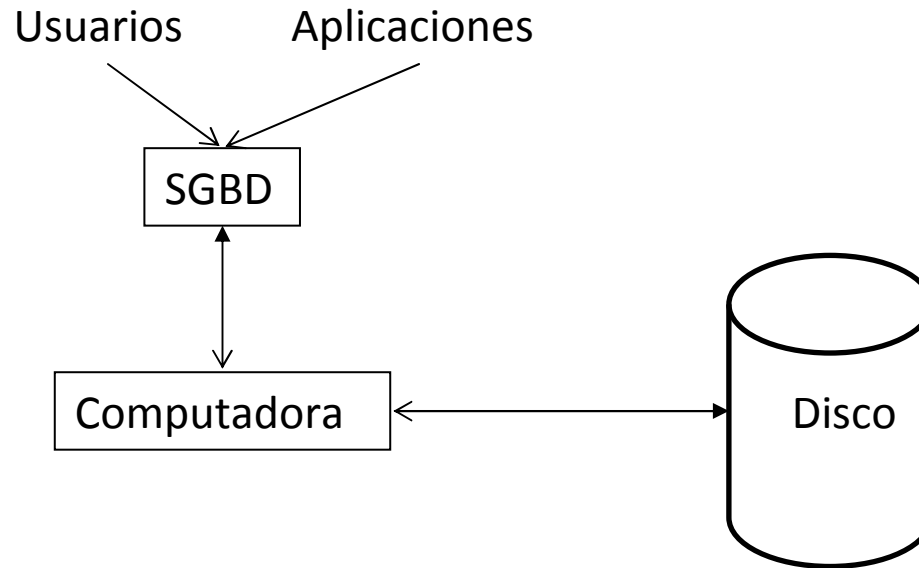
Dónde cada información de interés está representada, si es posible, solo una vez en la colección de datos.

Sistema Gestión de Bases de Datos es un sistema que permite gestionar solicitudes de usuarios y de aplicaciones de software en términos de consultas y transacciones para acceder a colecciones de datos (bases de datos) que, normalmente son:

Grande (mucho más grande que la disponibilidad de memoria principal).

Compartida (utilizada por distintas aplicaciones y usuarios).

Persistente (la vida de la información no está limitada a ejecuciones únicas de aplicaciones que utilizan dicha información).



Las solicitudes de los usuarios y de las aplicaciones son de dos tipos:

1. *Consultas*, obtienen información de los datos, existentes en la base de datos , que satisfacen un cierto criterio de selección. *No modifican el estado* de la base de datos.

2. *Transacciones*, que insertan, borran y actualizan datos en la base de datos. Por tanto, las transacciones *cambian el estado* de la base de datos.

Etapas del diseño

- » ¿qué pasos se siguen en el diseño de bases de datos?
- » existen una serie de diseños que se realizan en un cierto orden

- Etapas de diseño

- Análisis de requisitos

- *comprender* los datos a gestionar
 - necesidades del cliente: reuniones, discusiones, documentaciones, ...
 - etapa clave que puede ser muy costosa

...

– Diseño conceptual

- descripción de *alto nivel* de los datos y sus restricciones
- modelo que representa, organiza, *clarifica* la información
—habitualmente Entidad-Relación—
- *preciso*, que permita la traducción a un modelo específico del SGBD

– Diseño lógico

- esquema de la BD acorde al SGBD elegido
- (*relacional*) traducir esquema ER a esquema relacional

– Refinamiento de los esquemas

- reestructuración para garantizar propiedades importantes
—*normalización*—



– Diseño físico

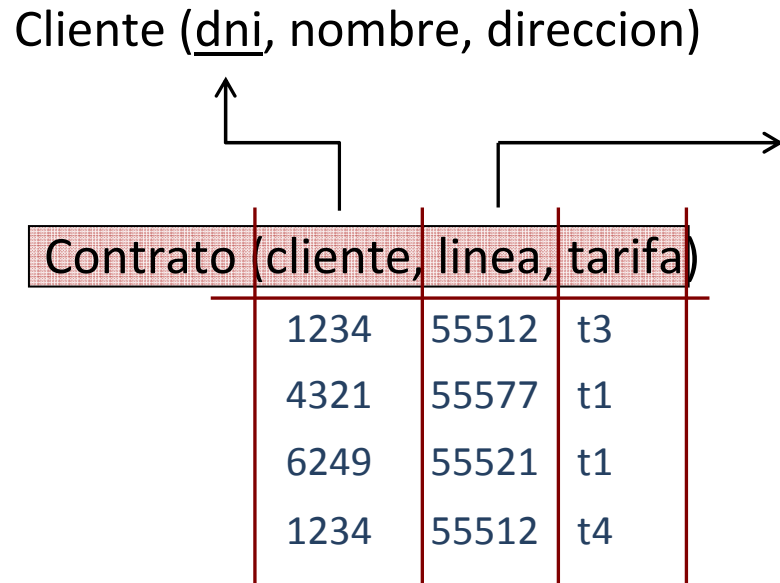
- mejora de rendimiento en base a cargas típicas
- idealmente no supone rediseño de etapas anteriores

– Diseño de aplicaciones y seguridad

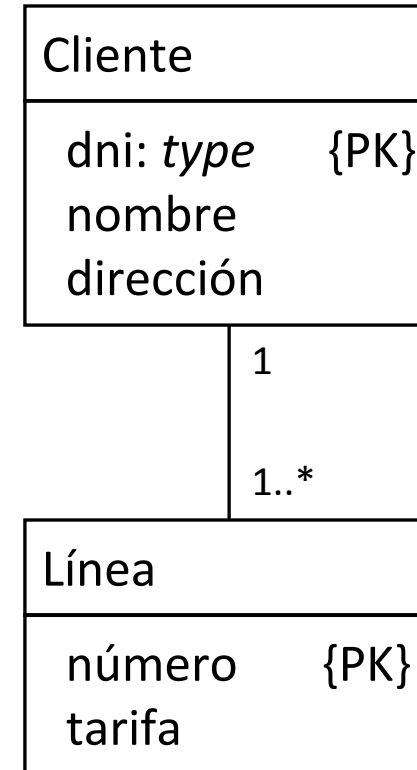
- procesos relacionados con las aplicaciones
- tareas y flujos de trabajo
- accesibilidad y seguridad

D. lógico (o conceptual) primero

» sobre el sistema de gestión telefónica



- ¿dos clientes pueden tener la misma línea?
- ¿una línea puede tener dos tarifas?
- ¿hay que poner más tablas? ¿cuáles serían?



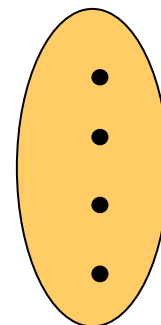
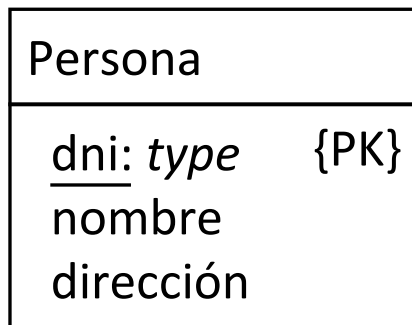
- ¿ER o UML?
- ER en notación UML

Diseño conceptual ER

- » descripción de cómo se *estructuran* los datos
- » varias alternativas para un mismo escenario

- **Entidades y conjuntos de entidad**

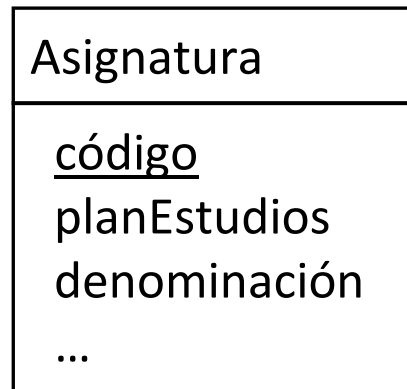
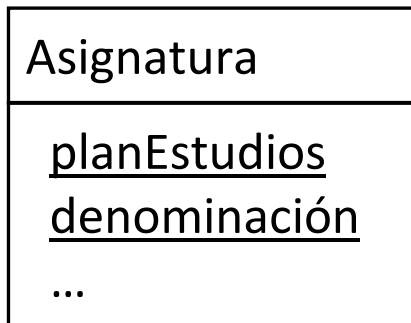
- objetos que engloban los datos de interés
- se describen como colecciones de entidades similares
- se describen mediante atributos y *propiedades* adicionales
- instancias como elementos de un conjunto



[123321,Santiago,'Plaza Mayor 1']

• Atributos (:dominios) y claves

- definición de datos para cada instancia
- el dominio (opcional) es el conjunto posible de valores; predefinidos o posibilidad de definición de tipos
- clave: conjunto mínimo de atributos que *identifican* a cada entidad (valor mínimo sin repetición)



- identificación más sencilla
- ¿cómo se asegura la no repetición plan + denominación?
- claves alternativas

Elementos del modelado ER

- **Decisiones de diseño**
 - Modelar un concepto como entidad o como atributo
 - Modelar un concepto como entidad o como relación
 - Identificación de relaciones: binarias o ternarias
 - Uso de la agregación
- **Restricciones en el modelo ER**
 - Los datos están acompañados de condiciones de validez
 - Algunas de estas restricciones no se pueden capturar en diagramas ER
 - Inclusión de restricciones en notas
 - Lenguajes de restricciones

Modelo relacional

Modelo Relacional: Introducción

Es un modelo sencillo y potente. Una BD es un conjunto de tablas con filas y columnas. Las tablas pueden estar relacionadas unas con otras.

La idea es antigua (Codd, CACM 1970) y es el modelo utilizado más ampliamente. Las razones de su uso se pueden resumir en:

- La representación de los datos es sencilla.

- Es muy fácil formular consultas en este modelo.

Ejemplos de SGBD de estas características:

- Ingres/Postgres, Informix (IBM), Sybase (SAP), MySQL (Oracle), SQL Server (Microsoft), Oracle, DB2 (IBM), ...

Modelo Relacional: Partes

Una Base de datos relacional es un conjunto de relaciones.

Una relación tiene dos partes:

- Esquema: nombre de la relación, nombre y tipo de cada columna (atributo, campo)
- Instancia: valores de la tabla (registro, tupla). Todas las tuplas son distintas.

Ejemplo: Películas

título	año	duración	género
Lo que el viento se llevó	1939	231	Drama
Star Wars	1977	124	C. Ficción
El mundo de Wayne	1992	95	Comedia

Modelo Relacional: Definiciones

Las columnas representan atributos (título, año, duración, género). Cada fila representa la información que considera el modelo para cada película. Se llama *tupla*. Una *tupla* tendrá un *componente* (valor) para cada atributo.

El nombre de una relación y el conjunto de atributos de la misma se llama *esquema*. Para la relación Películas el esquema es:

Películas (título, año, duración, género)

Los atributos son un conjunto y no una lista, aunque a veces se habla de orden estándar de los atributos.

El esquema para la relación Películas se puede representar por:

Películas (título: string, año: date, duración: int, género: string)

Se pueden listar los tres elementos de la relación en cualquier orden.

Modelo Relacional: Presentaciones de una relación

título	año	duración	género
Lo que el viento se llevó	1939	231	Drama
Star Wars	1977	124	C. Ficción
El mundo de Wayne	1992	95	Comedia

Distintas presentaciones de la relación Películas

año	género	título	duración
1939	Drama	Lo que el viento se llevó	231
1977	C. Ficción	Star Wars	124
1992	Comedia	El mundo de Wayne	95

género	título	año	duración
Drama	Lo que el viento se llevó	1939	231
C. Ficción	Star Wars	1977	124
Comedia	El mundo de Wayne	1992	95

Modelo Relacional: Creación de relaciones

- Creación de la relación/tabla, con sus atributos/campos
- Los tipos/dominios se especifican para cada campo
- El tipado es asegurado por el SGBD cuando se añaden o modifican tuplas/registros

- Es habitual la existencia de diversas tablas que, eventualmente, estarán relacionadas

```
CREATE TABLE Películas (  
    título    CHAR(50),  
    año      DATE,  
    duración INTEGER,  
    genero   CHAR(20)  
)
```

```
CREATE TABLE Directores (  
    título    CHAR(50),  
    director  CHAR(50)  
)
```

Nota: las tablas son un ejemplo, no necesariamente la mejor opción.

Modelo Relacional: Destrucción/ modificación

- La destrucción de una relación implica su borrado del esquema, y el borrado de todas sus tuplas

```
DROP TABLE Películas
```

- La modificación de un esquema se puede hacer añadiendo, borrando o modificando campos

- Para campos añadidos, todas las tuplas son extendidas con valor null es ese campo

```
ALTER TABLE Películas  
ADD COLUMN  
nacionalidad CHAR(20)  
)
```

Modelo Relacional: Destrucción/ modificación

- Se puede insertar una tupla en la relación

```
INSERT INTO Películas (título, año, duración, género)
VALUES ('El hombre tranquilo', 1951, 129, 'comedia')
```

- El borrado se realiza indicando la condición que cumplirán todas las tuplas a ser borradas

```
DELETE
FROM Películas E
WHERE E.título='Star Wars'
```

- La modificación se realiza igualmente con una condición de selección de tuplas

```
UPDATE Películas E
SET
E.duración=E.duración+1
WHERE E.título='El hombre tranquilo'
```

Restricciones de integridad

- **RI = condición de validez**
 - Condiciones que tienen que cumplirse para cualquier instancia de la base de datos
 - Se especifican cuando se define el esquema
 - Se comprueban cuando se modifican las relaciones (DBMS)
- **Instancia legal = satisface todas las RI**
- **Semántica del mundo real descrito**
- **Dominio, clave primaria, clave foránea**
- **También se soportan otras más generales**

Claves primarias y candidatas

- Clave (candidata)
 - Identificación única de cada tupla, por medio de un subconjunto mínimo de campos
- Condiciones
 - No existen dos tuplas con los mismos valores en todos los campos de la clave —implicación sobre los *null*
 - Ningún subconjunto de la clave es identificador único —superclave: {nia, nombre}
- Una candidata debe ser elegida como clave primaria (o principal)
 - ¿Quién hace esta elección?

Estudiante (nia, nombre, login, edad, notam)



- Las claves candidatas se especifican con **UNIQUE**, y la que es elegida como primaria con **PRIMARY KEY**

```
CREATE TABLE Matricula (  
    nia    INTEGER,  
    cod    CHAR(20),  
    nota   REAL,  
    PRIMARY KEY (nia,cod)  
)
```

```
CREATE TABLE Estudiante (  
    nia    INTEGER,  
    nombre CHAR(20),  
    login  CHAR(10),  
    edad   INTEGER,  
    notam  REAL,  
    UNIQUE (nombre,edad),  
    CONSTRAINT claveEst PRIMARY KEY (nia)  
)
```

nombrado de la restricción, lo que permite identificar errores

Modelo Relacional: Ejemplo

Relaciones correspondiente a los datos climatológicos de Segovia. La parte de año es un resultado elaborado no datos iniciales.

Mes	T	TM	Tm	R	H	DR	DN	DT	DF	DH	DD	I
Enero	4.3	8.2	0.3	38	74	6.9	-	0.0	3.9	14.5	4.5	124
Febrero	5.8	10.4	1.1	31	66	6.0	3.1	0.0	2.0	10.0	4.8	152
Marzo	8.6	13.9	3.2	30	59	5.9	1.7	0.1	1.3	6.1	5.4	203
Abril	9.7	15.1	4.2	44	59	8.0	1.5	0.8	0.6	4.2	3.4	213
Mayo	14.0	19.7	8.2	66	57	10.1	0.2	3.7	0.7	0.5	3.3	250
Junio	19.0	25.8	12.1	43	48	4.8	0.0	4.2	0.6	0.0	6.9	314
Julio	22.2	29.7	14.6	17	39	2.7	0.0	2.6	0.2	0.0	12.5	358
Agosto	22.1	29.4	14.8	20	40	3.0	0.0	3.8	0.1	0.0	10.4	328
Septiembre	17.7	24.0	11.4	28	50	4.7	0.0	2.1	0.4	0.0	5.8	246
Octubre	13.0	18.0	7.9	59	63	8.9	0.1	0.6	1.0	0.7	4.0	177
Noviembre	7.6	11.8	3.4	52	72	8.6	1.6	0.0	2.7	5.8	3.9	126
Diciembre	5.1	8.8	1.3	46	75	8.2	2.2	0.0	4.1	11.7	3.6	110
Año	12.4	17.9	6.9	479	59	78.6	-	18.1	17.6	53.1	-	-

T	Temperatura media mensual/anual (°C)
TM	Media mensual/anual de las temperaturas máximas diarias (°C)
Tm	Media mensual/anual de las temperaturas mínimas diarias (°C)
R	Precipitación mensual/anual media (mm)
H	Humedad relativa media (%)
DR	Número medio mensual/anual de días de precipitación superior o igual a 1 mm
DN	Número medio mensual/anual de días de nieve
DT	Número medio mensual/anual de días de tormenta
DF	Número medio mensual/anual de días de niebla
DH	Número medio mensual/anual de días de helada
DD	Número medio mensual/anual de días despejados
I	Número medio mensual/anual de horas de sol

Referencias

T. Connolly, C. Begg. *Sistemas de bases de datos*. 5a edición. Addison-Wesley. 2010.
Existe una edición de 2015.

A.Silberschatz, H. F. Korth, S. Sudarshan. *Fundamentos de bases de datos*. 6a edición.
McGraw-Hill. 2010. Material en: <http://codex.cs.yale.edu/avi/db-book/db6/slide-dir/>

[Elmasri02] R. Elmasri, S. Navathe. *Fundamentos de sistemas de bases de datos*.
3a edición. Addison-Wesley. 2002.