

BASE DE DATOS 1 – RESUMEN DE LA TEORIA

Modelo de datos

Provee una notación para describir los datos. Generalmente cuenta de estructura de los datos, y operaciones y restricciones sobre los datos.

Constituye la estructura subyacente de una base de datos.

- Clasificación:
 - o Modelos lógicos: modelo de datos de alto nivel que provee conceptos cercanos a la manera en la que los usuarios perciben los datos.
 - o Modelos físicos: modelo de datos de alto nivel que provee conceptos que describen detalles de como los datos son almacenados.
- Estos modelos difieren en los elementos que emplean para representar los datos y en expresividad.
- Modelos lógicos basados en objetos:
 - o Modelo de entidad y relaciones.
 - o Modelo orientado a objetos.
- Modelos lógicos basados en registros:
 - o Modelo relacional.

Modelo de entidades y relaciones

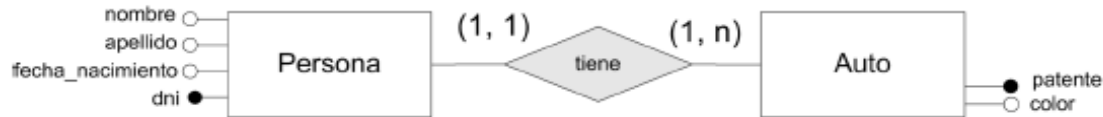
Es un modelo de datos lógico. Obtiene abstracción del mundo real, representa el significado de los datos y es independiente de los detalles de la implementación física.

Permite la comunicación entre personas.

- Elementos:
 - o Entidad: es una “cosa o concepto” que puede ser identificada y distinguible de otra “cosa o concepto”. Ejemplo: Juan con dni 1234567 – Auto modelo 2015 patente PRI
 - o Relación: es una asociación de entidades. Ejemplo: Juan con dni 1234567 *es_dueño_de* un auto modelo 2015 cuya patente es PRI.
 - o Atributos: representa información acerca de una entidad o una relación. Ejemplo: nombre, dni, modelo, patente.
 - Dominio de un atributo: conjunto de valores que puede tomar un atributo en particular. Ejemplo: nombre puede ser una cadena de máximo 50 letras del abecedario.
 - o Rol de una entidad en una relación: indica la función que tiene la entidad en la relación. Ejemplo: *esposa_de* Juan con dni 123456 *esta_casado* con Maria cuyo dni es 234567. Esta ultima tiene el rol de *esposa_de*
 - o Conjunto de entidades: deben ser del mismo tipo. Ejemplo: el conjunto de todas las personas que poseen un nombre y tienen dni puede llamarse PERSONA. El conjunto de todos los autos que poseen información del modelo y de la patente puede llamarse AUTO.
No necesariamente son disjuntos, por ejemplo, una entidad ALUMNO puede ser una entidad DOCENTE.
 - o Conjunto de relaciones: deben ser del mismo tipo. Ejemplo: ES_DUEÑO_DE es un conjunto de relaciones entre las entidades persona y auto.
- Restricciones:
 - o Cardinalidad: determina el numero de veces en el que puede participar una entidad en una relación. Indica dependencia, puede ser total o de existencia (participación obligatoria) o parcial (participación no obligatoria).
Tomamos un conjunto binario de relaciones R entre dos conjuntos A y B, la cardinalidad puede ser:
 - Uno a uno: una entidad de A puede estar asociada con a lo sumo una entidad de B y una entidad de B puede estar asociada con a lo sumo una entidad de A.
 - Uno a muchos: una entidad de A esta asociada con cualquier numero de entidades de B, pero una entidad de B esta asociada con a lo sumo una entidad de A.
 - Muchos a muchos: una entidad de A esta asociada con cualquier numero de entidades de B y una entidad de B esta asociada con cualquier numero de entidades en A.

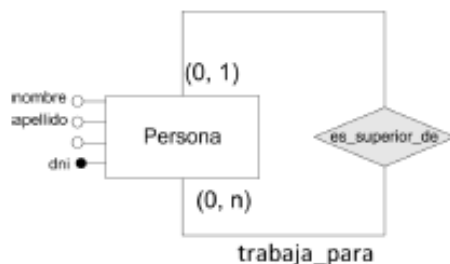
- Identificador o clave: restricción de unicidad del valor del atributo. Sirven para identificar de manera única a una entidad. Toda entidad posee al menos una posible clave o identificador. Puede ser simple o compuesto.
- Grado: representa el número máximo de veces que una entidad puede estar relacionada con otra.
(1, n) → grado n. (1, 1) → grado 1.
- Diagrama: representación gráfica de la estructura de los datos. La identidad se representa con un cuadrado, la relación con un rombo, los atributos como un chupetín y la cardinalidad entre paréntesis.

Ejemplo:



Este diagrama se lee: “una persona tiene al menos un auto y a lo sumo n. Y un auto es poseído por una única persona.”

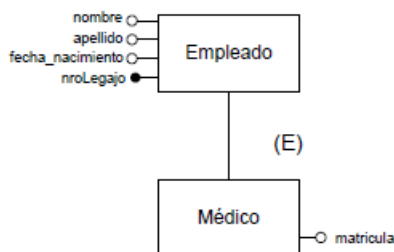
Ejemplo de Rol:



Y se lee: “una persona es superior de cero o muchas personas. Una persona trabaja para a lo sumo una persona.”

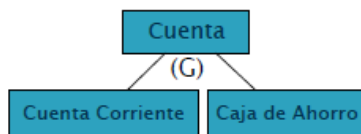
- Especialización: es el resultado de tomar un subconjunto de entidades de un nivel para formar un conjunto de entidades de nivel más bajo.

Ejemplo: tenemos empleados de un hospital. De los médicos nos interesa su matrícula. Puede haber empleados que no son médicos.



- Generalización: es el resultado de tomar uno o mas conjuntos de entidades (de nivel mas bajo) y producir un conjunto de entidades de un nivel más alto.

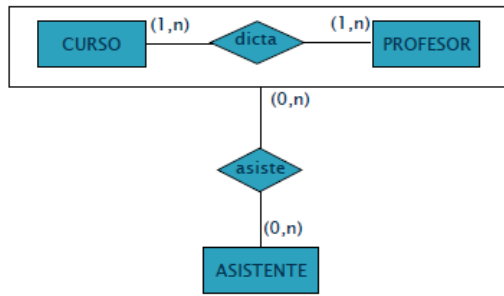
Ejemplo: distintos tipos de cuenta: caja de ahorro y cuenta corriente, pero ambas son cuentas.



- Agregación: es un mecanismo de abstracción en el cual una relación binaria (junto a las dos entidades relacionadas) se trata como entidad de alto nivel.

La cardinalidad máxima para cada entidad de la relación siempre debe ser mayor a 1, o sea n.

Ejemplo: un profesor puede dictar uno o varios cursos. Una vez asignado un profesor a un curso es posible que se registren asistentes a dicha asignación.



Modelo relacional

Representa los datos como tablas bidimensionales llamadas relaciones y cumple que:

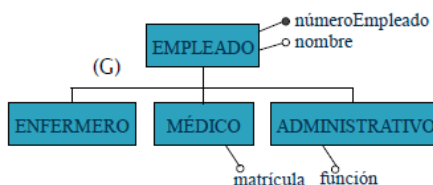
- No hay tuplas repetidas (filas).
- El orden de las filas no es significativo.
- El orden de las columnas no es significativo.
- Todos los valores de la tabla son atómicos.

El esquema está formado por el nombre de una relación y su conjunto de atributos. Ejemplo: Persona (dni,edad,nombre)

- Tuplas: son filas de una relación (excepto sus encabezados). Posee un solo componente para cada atributo de la relación. Ejemplo: (123456, 54, Juan) es una tupla con tres componentes.
- Dominio: cada componente de cada tupla debe ser atómica, es decir, debe ser un tipo elemental (no puede ser una lista, un registro, etc.).
- Clave de una relación: un conjunto de atributos conforma una clave en la relación cuando a dicho conjunto no se le permite tomar dos valores iguales en todos los atributos de la clave. Ejemplo: Persona (dni, edad, nombre).

Conversión del modelo E/R al modelo relacional:

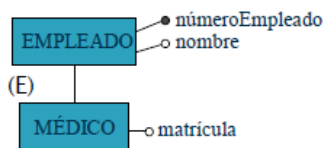
- Convertir cada conjunto de entidades en una relación (con igual nombre) con el mismo conjunto de atributos.
- Convertir cada relación del modelo de entidades y relaciones en una relación (del modelo relacional), de igual nombre.
 - Para cada entidad involucrada en la relación, se toma el o los atributos claves como parte del esquema de la relación (del modelo relacional).
 - Si la relación (del modelo E/R) posee atributos, estos también forman parte del esquema de la relación.
 - Si una entidad está involucrada más de una vez en una relación, con diferentes roles, se renombrará el atributo para evitar nombres duplicados, adoptando el nombre del rol de la entidad en la relación.
- Generalización: hay tres opciones para realizar el pasaje a tablas.



- Una tabla para el conjunto de entidades nivel más alto.
 - Empleado (numeroEmpleado,nombre,tipoEmpleado,matricula,función)
- Una tabla para cada conjunto de entidades del nivel más bajo.
 - Enfermero (numeroEmpleado,nombre)
 - Medico (numeroEmpleado,nombre,matricula)
 - Administrativo (numeroEmpleado,nombre,función)
- Una tabla para el conjunto de entidades de nivel mas alto y una tabla para cada conjunto de entidades del nivel más bajo.
 - Empleado (numeroEmpleado,nombre)
 - Enfermero (numeroEmpleado)
 - Medico (numeroEmpleado,matricula)

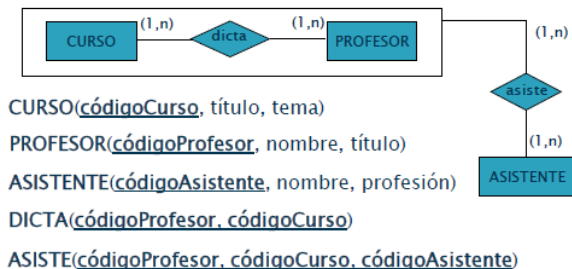
- Administrativo (numeroEmpleado,función)

- Especialización: hay dos opciones para realizar el pasaje a tablas.



- Una tabla para el conjunto de entidades de nivel más alto.
 - Empleado (numeroEmpleado,nombre,tipoEmpleado,matricula)
- Una tabla para el conjunto de entidades de nivel mas alto, y una tabla para cada conjunto de entidades del nivel más bajo.
 - Empleado (numeroEmpleado,nombre)
 - Medico (numeroEmpleado,matricula)

- Agregación:



Algebra Relacional

Es un lenguaje de consulta, procedimental, que incluye operaciones fundamentales, adicionales y de manipulación de datos. Estas operaciones se pueden usar aisladas o combinadas (expresiones), y permiten resolver consultas complejas. Se usan paréntesis cuando es necesario agrupar operaciones y tiene notación lineal.

- Operaciones fundamentales: son suficientes para expresar cualquier consulta en algebra relacional.
 - Selección: operación unaria. Requiere una condición booleana (and, or y not). El resultado es una relación con un subconjunto “horizontal” de la relación dada.

Ingenieros

E#	Nombre	Edad
320	José	34
322	Rosa	37
323	María	25

$\sigma_{\text{edad} \geq 30}(\text{Ingenieros})$

E#	Nombre	Edad
320	José	34
322	Rosa	37

- Proyección: operación unaria. Dada una lista de atributos produce un corte “vertical” de la relación.

Ingenieros

E#	Nombre	Edad
320	José	34
322	Rosa	37
323	María	25

$\Pi_{\text{nombre, edad}}(\text{Ingenieros})$

Nombre	Edad
José	34
Rosa	37
María	25

- Producto Cartesiano (X): operación binaria (A X B). El resultado es una relación que incluye todas las tuplas posibles que se obtienen concatenando cada tupla de A con cada una de las tuplas de B.

Ingenieros		
E#	Nombre	D#
320	José	D1
322	Rosa	D3

Proyectos	
Proyecto	Tiempo
RX338A	21
PY254Z	32

Ingenieros \bowtie Proyectos

E#	Nombre	D#	Proyecto	Tiempo
320	José	D1	RX338A	21
320	José	D1	PY254Z	32
322	Rosa	D3	RX338A	21
322	Rosa	D3	PY254Z	32

- Renombre (p): operación unaria ($p \times R$). El resultado es la relación R con nombre X.

Ingenieros		
E#	Nombre	Edad
320	José	34
322	Rosa	37
323	María	25

ρ PROFESIONALES Ingenieros

PROFESIONALES		
E#	Nombre	Edad
320	José	34
322	Rosa	37
323	María	25

- Unión (U): operación binaria ($A \cup B$). El resultado es una relación en la que se agrega a la relación A los elementos (no repetidos) de la relación B. Es necesario que las relaciones A y B sean de <unión compatible>: relaciones con igual aridad (igual número de atributos) y el dominio del i-esimo atributo de ambas relaciones debe ser el mismo.

Ingenieros		
E#	Nombre	Edad
320	José	34
322	Rosa	37
323	María	25

Jefes		
E#	Nombre	Edad
320	José	34
421	Jorge	48

Ingenieros \cup Jefes

E#	Nombre	Edad
320	José	34
322	Rosa	37
323	María	25
421	Jorge	48

- Diferencia (-): operación binaria ($A - B$). El resultado es una relación donde están los elementos que pertenecen a A y no pertenecen a B. Es necesario que las relaciones A y B sean de <unión compatible>: relaciones con igual aridad (igual número de atributos) y el dominio del i-esimo atributo de ambas relaciones debe ser el mismo.

Ingenieros		
E#	Nombre	Edad
320	José	34
322	Rosa	37
323	María	25

Jefes		
E#	Nombre	Edad
320	José	34
421	Jorge	48

Ingenieros $-$ Jefes

E#	Nombre	Edad
322	Rosa	37
323	María	25

- Operaciones adicionales: no agregan potencia al algebra, simplifican consultas. Son rescribibles en termino de operaciones fundamentales.

- Intersección: operación binaria. El resultado es una relación con aquellas tuplas que pertenecen a ambas relaciones (al mismo tiempo). Es necesario que las relaciones A y B sean de <unión compatible>. La consulta es equivalente a $R - (R - S)$.

Ingenieros		
E#	Nombre	Edad
320	José	34
322	Rosa	37
323	Maria	25

Jefes		
E#	Nombre	Edad
320	José	34
421	Jorge	48

Ingenieros \cap Jefes

E#	Nombre	Edad
320	José	34

- **Producto Theta:** operación binaria ($R \bowtie S$). Genera una nueva relación con las tuplas resultantes de aplicar una operación de selección con la condición indicada por θ sobre el resultado de un producto cartesiano. La condición (θ) se indica como una expresión booleana de términos (se pueden usar conectores lógicos entre las condiciones). La consulta es equivalente a selección o ($R \bowtie S$).

Ingenieros		
E#	Nombre	Edad
320	José	34
322	Rosa	37
323	Maria	25

Jefes		
E#	Nombre	Edad
320	José	34
421	Jorge	48

Ingenieros \bowtie Ingenieros.edad = Jefes.edad Jefes

Ingenieros.E#	Ingenieros.Nombre	Ingenieros.Edad	Jefes.E#	Jefes.Nombre	Jefes.Edad
320	José	34	320	José	34

- **Producto Natural (\bowtie):** operación binaria ($R \bowtie S$). Genera una nueva relación con las tuplas resultantes de aplicar una operación de selección con la condición indicada sobre el resultado de un producto cartesiano. La condición se indica como una expresión booleana de términos (se pueden usar conectores lógicos entre las condiciones). La consulta es equivalente a proyección (selección ($R \bowtie S$)).

Postulantes			Administrativos			
Nombre	Edad	DNI	Nombre	Edad	Domicilio	DNI
Paula	19	29235142	Martina	22	1 y 50	35215415
Martina	22	35215415	Paula	19	8 y 49	29899632
Joaquín	28	28152478	Pablo	32	26 y 50	20125789

Postulantes \bowtie Administrativos

Nombre	Edad	DNI	Domicilio
Martina	22	35215415	1 y 50

- **División (\div):** operación binaria ($R \div S$). R dividendo, S divisor. Los atributos del divisor S deben ser un subconjunto de los atributos de la relación R con igual dominio. La relación resultante de la división, llamémosla T , posee tuplas t tal que: los valores de t deben aparecer en R en combinación con todas las tuplas de S .

Lugar_Trabajo			Sucursales_Vip	
Nombre	Sucursal	Color	Sucursal	Color
Juan	Sucursal1	Rojo	Sucursal1	Rojo
Pedro	Sucursal1	Verde	Sucursal2	Azul
Juan	Sucursal2	Azul	Sucursal1	Verde
María	Sucursal 1	Rojo		
Juan	Sucursal3	Violeta		
Pedro	Sucursal1	Rojo		
Pedro	Sucursal2	Azul		
Juan	Sucursal1	Verde		

Lugar_Trabajo \div Sucursales_Vip

Nombre
Juan
Pedro

- **Asignación (\leftarrow):** es una forma conveniente de expresar operaciones complejas (modularidad). El resultado de una operación se asigna temporalmente a una variable. La variable a la cual se asigna el resultado de una operación se puede usar en otras operaciones.

Lugar_Trabajo		Sucursales_Vip		Empleado	
Nombre	Sucursal	Sucursal		Nombre	Domicilio
Juan	Sucursal1	Sucursal1		Juan	1 y 50
Pedro	Sucursal1	Sucursal2		Pedro	120 y 43
Juan	Sucursal2			María	150 y 62
María	Sucursal 1				
Juan	Sucursal3				

Empleados_Vip ← Lugar_Trabajo % Sucursales_Vip

Nombre
Juan

Empleados_Vip |X| Empleado

Nombre	Domicilio
Juan	1 y 50

- Operaciones de manipulación: se expresan usando la operación de asignación. Modifican la cantidad o los valores de las tuplas de una relación.

- **Inserción (U)**: la o las tuplas a insertar deben ser compatibles con la relación. $R \leftarrow R \cup E$, donde R es la relación donde se insertarán los resultados de la expresión E.

Empleado	
Nombre	Domicilio
Juan	1 y 50
Pedro	120 y 43
María	150 y 62

Empleado ← Empleado \cup {"Joaquín", "4 y 497"}

Nombre	Domicilio
Juan	1 y 50
Pedro	120 y 43
María	150 y 62
Joaquín	4 y 497

- **Eliminación (-)**: la o las tuplas a eliminar deben ser compatibles con la relación. $R \leftarrow R - E$, donde R es la relación donde se insertarán los resultados de la expresión E.

Empleado	
Nombre	Domicilio
Juan	1 y 50
Pedro	120 y 43
María	150 y 62
Joaquín	4 y 497
Martina	1 y 32

Empleado ← Empleado - {"Joaquín", "4 y 497"}, {"Martina", "1 y 32"}

Nombre	Domicilio
Juan	1 y 50
Pedro	120 y 43
María	150 y 62

- **Actualización**: permite actualizar un valor particular de una tupla.

Empleado		
Nombre	Domicilio	Salario
Juan	1 y 50	10200
Pedro	120 y 43	15000
María	150 y 62	22000

δ salario ← salario * 1.2 (Empleado)

Nombre	Domicilio	Salario
Juan	1 y 50	12240
Pedro	120 y 43	18000
María	150 y 62	26400