



# **Preparador Informática**

[www.preparadorinformatica.com](http://www.preparadorinformatica.com)

**TEMA 38. INFORMÁTICA**

**TEMA 37. S.A.I.**

**MODELO DE DATOS RELACIONAL.  
ESTRUCTURA. OPERACIONES.  
ÁLGEBRA RELACIONAL.**

## 1. INTRODUCCIÓN

Hoy en día el mayor activo de las organizaciones son los datos y su gestión eficaz y segura. Por ello, si analizamos la mayoría de los ámbitos de actividad, nos encontramos que la utilización de las bases de datos está ampliamente extendida (al registrarse en una web, al acudir a la consulta médica, al consultar el catálogo de productos de una tienda online, etc.). Las bases de datos y los datos contenidos en ellas, son imprescindibles para llevar a cabo multitud de acciones.

La clasificación tradicional de las bases de datos establece tres modelos de bases de datos: jerárquico, en red y relacional. En la actualidad el modelo de bases de datos más extendido es el relacional. Aunque, hay que tener en cuenta que dos de sus variantes (modelo de bases de datos distribuidas y orientadas a objetos) son las que se más se están utilizando en los últimos tiempos.

En el presente tema nos centraremos en analizar el **modelo de datos relacional**, describiendo su estructura, operaciones y álgebra relacional detallando los operadores básicos sobre los que se construye este último.

## 2. MODELO DE DATOS RELACIONAL

### 2.1. MODELO DE DATOS

En informática, un **modelo de datos** es un lenguaje utilizado para la descripción de una base de datos. Con este lenguaje vamos a poder describir las estructuras de los datos (tipos de datos y relaciones entre ellos), las restricciones de integridad (condiciones que deben cumplir los datos, según las necesidades de nuestro modelo basado en la realidad) y las operaciones de manipulación de los datos (insertado, borrado, modificación de datos). Es importante distinguir entre modelo de datos y esquema. Para clasificar los modelos debemos pensar en el nivel de abstracción, es decir, en lo alejado que esté del mundo real:

- Los modelos de datos **conceptuales** son aquellos que describen las estructuras de datos y restricciones de integridad. Se utilizan durante la etapa de análisis de un problema dado, y están orientados a representar los elementos que intervienen y sus relaciones. Ej: Modelo Entidad-Relación.

Ejemplo:



Para manipular la información utilizamos un lenguaje relacional, actualmente se cuenta con dos lenguajes formales el Álgebra relacional y el Cálculo relacional. El Álgebra relacional permite describir la forma de realizar una consulta, en cambio, el Cálculo relacional solamente indica lo que se desea devolver.

Es el modelo más utilizado en la actualidad para modelar problemas reales y administrar datos dinámicamente.

## Preparador Informática

### 2.2.1. VENTAJAS

- Provee herramientas que garantizan evitar la duplicidad de registros.
- Garantiza la integridad referencial, así, al eliminar un registro elimina todos los registros relacionados dependientes.
- Favorece la normalización por ser más comprensible y aplicable.

### 2.2.2. INCONVENIENTES

- Presentan deficiencias con datos gráficos, multimedia, CAD y sistemas de información geográfica.
- No se manipulan de forma eficiente los bloques de texto como tipo de dato.

A partir de ahora, nosotros veremos una relación como una tabla con filas y columnas que nos sirve para almacenar los datos de una entidad. Por tanto, en cada relación podemos distinguir su nombre, un conjunto de columnas, denominadas atributos, que representan propiedades de la tabla y que también estarán caracterizadas por su nombre, y un conjunto de filas llamadas tuplas, que contienen los valores que toma cada uno de los atributos para cada elemento de la relación.

Hay dos conceptos fundamentales en el modelo relacional:

- **Cardinalidad:** Es el número de filas de la tabla.
- **Grado:** Es el número de atributos de la tabla. Siempre todas las tuplas tienen el mismo número de atributos.

### 3.2. DOMINIO Y ATRIBUTO

**Dominio** es un conjunto finito de valores homogéneos y atómicos, caracterizados por un nombre. Homogéneos, porque son todos del mismo tipo, y atómicos porque son indivisibles en lo que al modelo se refiere. Todo dominio debe tener un nombre por el cual nos referimos a él y un tipo de datos.

Un **atributo** es el papel que tiene un determinado dominio en una relación.

El dominio contiene todos los posibles valores que puede tomar un atributo y es estático. Los dominios representan un papel importante y característico en ciertas operaciones.

También existen los denominados dominios compuestos que son combinaciones de dominios simples a las que se pueden aplicar ciertas restricciones de integridad.

### 3.3. CLAVES

Hay distintos tipos de claves:

- a) Clave **candidata**: Es un conjunto no vacío de atributos que identifican unívoca y mínimamente cada tupla.
- b) Clave **primaria**: Es la que el usuario escoge entre las claves candidatas. Debe ser única y mínima, no nula y no ambigua.



- Operación con transmisión en cascada: El borrado o la modificación de una fila de la tabla que contiene la clave primaria lleva consigo el borrado o la modificación en cascada de las filas de la tabla que referencia cuya clave ajena coincide con el valor de la clave primaria de la tabla referenciada.
- Operación con puesta a nulos: El borrado o la modificación de una fila de la tabla que contiene la clave primaria lleva consigo la puesta a nulos de los valores de la clave ajena de las filas de la tabla que referencia cuya clave ajena coincide con el valor de la clave primaria de la tabla referenciada.

### 3.5. ÍNDICES

En las bases de datos, cada tabla se divide internamente en páginas de datos, y se define el índice a través de un campo (o campos) y es a partir de este campo desde donde se busca.

Un **índice** es una estructura de datos que permite acceder a diferentes filas de una misma tabla a través de un campo o campos. Esto permite un acceso mucho más rápido a los datos. Los índices son útiles cuando se realizan consultas frecuentes a un rango de filas o una fila de una tabla.

Los cambios en los datos de las tablas (agregar, actualizar o borrar filas) son incorporados automáticamente a los índices con transparencia total.

Los índices son independientes, lógica y físicamente de los datos, es por eso que pueden ser creados y eliminados en cualquier momento, sin afectar a las tablas ni a otros índices. No hay un límite de columnas a indexar, si quisiéramos podríamos crear un índice para cada columna, pero no sería operativo. Normalmente tiene sentido crear índices para ciertas columnas ya que agilizan las operaciones de búsqueda de base de datos grandes.

Al crear índices, las operaciones de modificar o agregar datos se ralentizan, ya que al realizarlas es necesario actualizar tanto la tabla como el índice. Si se elimina un índice, el acceso a datos puede ser más lento a partir de ese momento.

### 5.1. OPERADORES PRIMITIVOS:

- a) **Unión de conjuntos:** La unión de dos relaciones R y S, denotada por  $R \cup S$ , es el conjunto de tuplas que pertenece a R, a S ó a ambas. Sólo se aplica este operador a relaciones del mismo grado y definidas sobre los mismos atributos.
- b) **Diferencia de conjuntos:** La diferencia de dos relaciones R y S, denotada por  $R - S$ , es el conjunto de tuplas de R que no pertenecen a S. Las mismas restricciones del caso anterior se verificarán sobre R y S.
- c) **Producto cartesiano:** Sean R y S dos relaciones de grados m y n respectivamente. El producto cartesiano  $R \times S$  es una relación de grado  $m+n$ , constituida por todas las posibles tuplas, en que los m primeros elementos constituyen una tupla de R, y los n últimos una tupla de S.
- d) **Proyección:** La proyección  $\pi_x(R)$ , donde R es una relación definida sobre J (esquema o conjunto de atributos posibles) y x está incluido o coincide con J, es una relación construida por las columnas de R correspondientes a los atributos de x.
- e) **Selección:** Sea F una fórmula que involucra:
- Operandos constantes, y que referencian a atributos de la relación implicada R.
  - Operadores aritméticos de comparación:  $<, >, =, <=, >=, <>$ .
  - Operadores lógicos: AND, OR.

Entonces la selección  $\sigma_F(R)$  es el conjunto de tuplas de R tales que una vez sustituida en la fórmula F, las ocurrencias de los atributos que en ellas se referencian, resulta verdadera.

### 5.2. OPERADORES DERIVADOS:

- a) **Intersección:** Es el inverso a la resta, y se resuelve por las tuplas comunes a las dos relaciones que se intersectan
- $$R \cap S = (R - (R - S))$$
- b) **División o cociente:** Sean R y S relaciones de grado r y s respectivamente, donde  $r > s$  y  $S \neq \emptyset$ . Entonces el cociente  $R:S$  es el conjunto de tuplas t de grado  $(r - s)$  tales que para toda tupla u de S, la tupla t está en R.