

# Conceptos básicos de integridad referencial.

## Introducción

La **integridad referencial** es un sistema de **reglas** que utilizan la mayoría de las bases de datos relacionales para **asegurarse que los registros de tablas relacionadas son válidos** y que no se borren o cambien datos relacionados de forma accidental produciendo errores de integridad.

Primero repasemos un poco los tipos de relaciones.

## Tipos de relaciones.

Entre dos tablas de cualquier base de datos relacional pueden haber dos tipos de relaciones, relaciones uno a uno y relaciones uno a muchos:

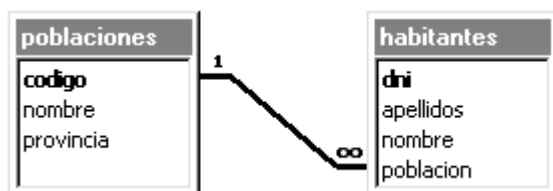
● **Relación Uno a Uno:** Cuando **un registro de una tabla sólo puede estar relacionado con un único registro de la otra tabla y viceversa**.

Por ejemplo: tenemos dos tablas una de profesores y otra de departamentos y queremos saber qué profesor es jefe de qué departamento, tenemos una relación uno a uno entre las dos tablas ya que un departamento tiene un solo jefe y un profesor puede ser jefe de un solo departamento.

● **Relación Uno a Varios:** Cuando **un registro de una tabla** (tabla secundaria) **sólo puede estar relacionado con un único registro de la otra tabla** (tabla principal) **y un registro de la tabla principal puede tener más de un registro relacionado en la tabla secundaria**, en este caso se suele hacer referencia a la tabla principal como tabla 'padre' y a la tabla secundaria como tabla 'hijo', entonces la regla se convierte en 'un padre puede tener varios hijos pero un hijo solo tiene un padre (regla más fácil de recordar).

Por ejemplo: tenemos dos tablas una con los datos de diferentes poblaciones y otra con los habitantes, una población puede tener más de un habitante, pero un habitante pertenecerá (estará empadronado) en una única población. En este caso la tabla principal será la de poblaciones y la tabla secundaria será la de habitantes. Una población puede tener varios habitantes pero un habitante pertenece a una sola población. Esta relación se representa incluyendo en la tabla 'hijo' una columna que se corresponde con la clave principal de la tabla 'padre', esta columna la denominamos clave foránea (o clave ajena o clave externa).

Una clave foránea es pues un campo de una tabla que contiene una referencia a un registro de otra tabla. Siguiendo nuestro ejemplo en la tabla habitantes tenemos una columna población que contiene el código de la población en la que está empadronado el habitante, esta columna es clave ajena de la tabla habitantes, y en la tabla poblaciones tenemos una columna código de población clave principal de la tabla.

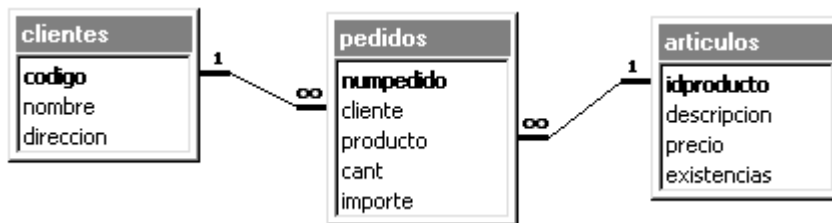


● **Relación Varios a Varios:** Cuando **un registro de una tabla puede estar relacionado con más de un registro de la otra tabla y viceversa**. En este caso las dos tablas no pueden estar relacionadas directamente, se tiene que añadir una tabla entre las dos que incluya los pares de valores relacionados entre sí.

Por ejemplo: tenemos dos tablas una con los datos de *clientes* y otra con los *artículos*

que se venden en la empresa, un cliente podrá realizar un pedido con varios artículos, y un artículo podrá ser vendido a más de un cliente.

No se puede definir entre *clientes* y *artículos*, hace falta otra tabla (por ejemplo una tabla de pedidos) relacionada con clientes y con artículos. La tabla pedidos estará relacionada con cliente por una relación uno a muchos y también estará relacionada con artículos por un relación uno a muchos.



## Integridad referencial

Cuando se define una columna como clave foránea, las filas de la tabla pueden contener en esa columna o bien el valor nulo (ningún valor), o bien un valor que existe en la otra tabla, un error sería asignar a un habitante una población que no está en la tabla de poblaciones. Eso es lo que se denomina **integridad referencial** y consiste en que **los datos que referencian otros (claves foráneas) deben ser correctos**. La **integridad referencial** hace que el sistema gestor de la base de datos se asegure de que no hay en las claves foráneas valores que no estén en la tabla principal.

**La integridad referencial se activa en cuanto creamos una clave foránea** y a partir de ese momento se comprueba cada vez que se modifiquen datos que puedan alterarla.

### ¿Cuándo se pueden producir errores en los datos?

● **Cuando insertamos una nueva fila en la tabla secundaria y el valor de la clave foránea no existe en la tabla principal.** Insertamos un nuevo habitante y en la columna *poblacion* escribimos un código de población que no está en la tabla de poblaciones (una población que no existe).

● **Cuando modificamos el valor de la clave principal de un registro que tiene 'hijos'.** Modificamos el código de Valencia, sustituimos el valor que tenía (1) por un nuevo valor (10), si Valencia tenía habitantes asignados, qué pasa con esos habitantes, no pueden seguir teniendo el código de población 1 porque la población 1 ya no existe, en este caso hay dos alternativas, no dejar cambiar el código de Valencia o bien cambiar el código de población de todos los habitantes de Valencia y asignarles el código 10.

● **Cuando modificamos el valor de la clave foránea, el nuevo valor debe existir en la tabla principal.** Por ejemplo cambiamos la población de un habitante, tenía asignada la población 1 (porque estaba empadronado en valencia) y ahora se le asigna la población 2 porque cambia de lugar de residencia. La población 2 debe existir en la tabla de poblaciones.

● **Cuando queremos borrar una fila de la tabla principal y ese registro tiene 'hijos'.** Por ejemplo queremos borrar la población 1 (Valencia). Si existen habitantes asignados a la población 1, estos no se pueden quedar con el valor 1 en la columna población porque tendrían asignada una población que no existe. En este caso tenemos dos alternativas, no dejar borrar la población 1 de la tabla de poblaciones, o bien borrarla y poner a valor nulo el campo población de todos sus 'hijos'.

Asociada a la integridad referencial están los conceptos de actualizar los registros en cascada y eliminar registros en cascada.

## Actualización y borrado en cascada

El actualizar y/o eliminar registros en cascada, son opciones que se definen cuando definimos la clave foránea y que indican al sistema gestor qué hacer en los casos comentados en el punto anterior

### ● Actualizar registros en cascada:

Esta opción le indica al sistema gestor de la base de datos que **cuando se cambie un valor del campo clave de la tabla principal**, automáticamente deberá **cambiar el valor de la clave foránea de los registros relacionados en la tabla secundaria**.

Por ejemplo, si cambiamos en la tabla de poblaciones (la tabla principal) el valor 1 por el valor 10 en el campo código (la clave principal), automáticamente se actualizan todos los habitantes (en la tabla secundaria) que tienen el valor 1 en el campo población (en la clave ajena) dejando 10 en vez de 1.

Si no se tiene definida esta opción, no se puede cambiar los valores de la clave principal de la tabla principal. En este caso, si intentamos cambiar el valor 1 del código de la tabla de poblaciones, no se produce el cambio y el sistema nos devuelve un error o un mensaje indicándonos que los registros no se han podido modificar por infracciones de clave.

### ● Eliminar registros en cascada:

Esta opción le indica al sistema gestor de la base de datos que **cuando se elimina un registro de la tabla principal** automáticamente **se borren también los registros relacionados en la tabla secundaria**.

Por ejemplo: Si borramos la población Onteniente en la tabla de poblaciones, automáticamente todos los habitantes de Onteniente se borrarán de la tabla de habitantes.