



Universidad
Francisco de Vitoria
UFV Madrid

Fundamentos de la ingeniería informática

Ingeniería de sistemas industriales

Curso 2019-2020

Bases de Datos

1. Base de Datos

Una base de datos es una colección de datos organizados y estructurados según un determinado modelo de información que refleja no sólo los datos en sí mismos, sino también las relaciones que existen entre ellos.

Físicamente una base de datos es un conjunto de tablas relacionadas físicamente (vinculadas) entre sí.

2. Sistemas de gestión

Los SGDB (Sistemas de gestión de bases de datos) son un tipo de software exclusivamente dedicado a tratar con bases de datos

Los más populares son:

- Microsoft SQL Server
- Oracle
- PostgreSQL
- MySQL
- SQLite
- Access

Todos ellos ofrecen en mayor o menor grado, de forma más o menos sofisticada, prestaciones en:

- Control de la integridad de los datos.
Asegura en cierto modo que los datos son consistentes.
- Gestión de la concurrencia
Permite que varios “usuarios” operen sobre la base de datos de forma simultánea.
- Recuperación de información
Métodos para recuperar información a partir de los datos almacenados (consultas)
- Interfaz de usuario
Un entorno que permita operar sobre la base de forma sencilla.
- Lenguaje para la operación.

Casi todos los nombres de los SGDB hacen referencia a SQL y eso es porque todos ellos utilizan como lenguaje de operación el llamado SQL, aunque no todos comparten la versión de SQL que implementan que varían en ligeros detalles que afectan a operaciones menos habituales.

3. SQL

Structured Query Language (Lenguaje de consulta estructurada) es un lenguaje de dominio específico utilizado en programación, diseñado para administrar, y recuperar información de sistemas de gestión de bases de datos relacionales. Una de sus principales características es el manejo del álgebra

y el cálculo relacional para efectuar consultas con el fin de recuperar, de forma sencilla, información de bases de datos, así como realizar cambios en ellas.

El lenguaje permite principalmente acciones (consultas) de:

- Creación (de tablas y bases de datos)
 - Modificación de tablas.
 - Eliminación de tablas
 - Vaciado de tablas
-
- Inserción (creación de nuevos registros).
 - Actualización (modificación de registros existentes).
 - Borrado de registros.
 - Selección (recuperación de información).

Nos centraremos en estas cuatro últimas, pues son las operaciones habituales, mientras que para la creación, modificación y eliminación de tablas y bases de datos los interfaces de usuario ofrecen servicios que resultan más fáciles de manejar que comandos SQL.

Una **consulta** es una operación que se realiza sobre la base de datos en base a la especificación de dicha operación realizada en el correspondiente lenguaje; en nuestro caso SQL

4. Diseño de bases de datos

A la hora de realizar el diseño y construcción de una base de datos se recorren cinco etapas:

1. Especificación:

Se utiliza el lenguaje natural (español) para describir las características y comportamiento esperado de la base de datos. Ello incluye la descripción de los datos almacenados y de la información que se pretende recuperar.

2. Diseño conceptual:

Se trata de trasladar la especificación a un modelo E-R compuesto por su diagrama E-R y los comentarios descriptivos necesarios.

3. Diseño físico:

Decidir las tablas, los campos, las claves, los índices, y las relaciones físicas.

4. Construcción:

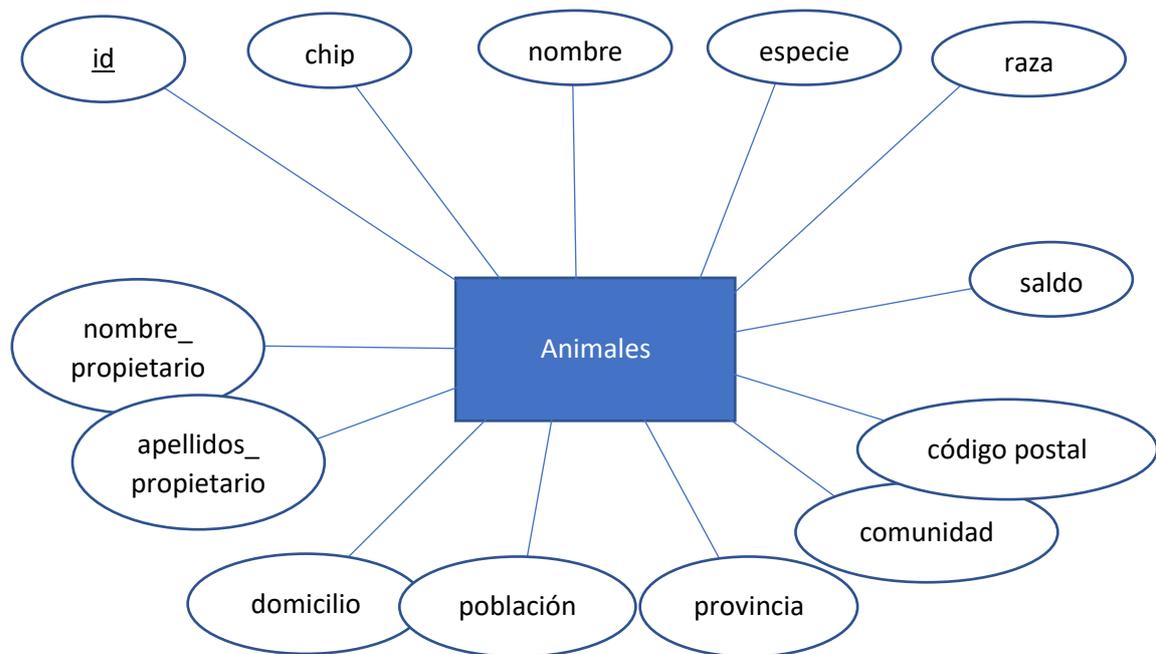
En el SGBD crear todos los elementos del diseño físico y poblar la base de datos.

5. Tablas

Una tabla es un conjunto de registros del mismo tipo. En una primera aproximación simplista, una tabla corresponde a la implementación física de un grupo de entidades del modelo E-R. Por ello las tablas permiten anotar los atributos de cada entidad elementos del registro llamados campos.

id	chip	nombre	especie	raza	nombre_propietario	apellidos_propietario	domicilio	poblacion	provincia	comunidad
1	982009100150632	JAK	PERRO	CHOW-CHOW	LEANDRO	LOPEZ VEIGUELA	AVDA. DE BETA	Guadalajara	Guadalajara	Castilla la
2	982009100141452	MARTINA	GATO	SIAMÉS	CONCHA	GOMEZ CARDIN	AVDA. DE FRAI	Mostoles	Madrid	C.A.Madri
3	982009100141414	BRUTO	PERRO	MESTIZO	NOELIA	PERNAUTE	COLLADO DEL	Ciudad Real	Ciudad Real	Castilla la
4	982009100141308	NAUS	GATO	EUROPEO DE P	OLGA	PERNAUTE	VENTISQUERO	Alcorcón	Madrid	C.A.Madri
5	982009100138878	BARTOK	GATO	PERSA	NIEVES	RABASSA GARCIA	CARDENAL HEF	Cuenca	Cuenca	Castilla la
6	982009100137490	PATY	PERRO	COCKER SPANI	JOSE	CLAVIJO QUESADA	NURIA,16 CHAI	Toledo	Toledo	Castilla la
7	982009100120235	NUQUI	GATO	PERSA	PILAR	BOUZAS VITURRO	NURIA, 5 CHAL	Alcobendas	Madrid	C.A.Madri
8	982009100116535	KLAUS	PERRO	PASTOR ALEM/	DAVID	PALOMO RUIZ	NURIA, 91 ESC.	Las Rozas	Madrid	C.A.Madri
9	982009100114324	POPI	PERRO	PASTOR ALEM/	MERCEDES	RODRIGUEZ ANDIA-ROMER	ORTALEZA, 10€	Mostoles	Madrid	C.A.Madri
10	982009100039117	DORA	GATO	EUROPEO DE P	ALFONSO	GOMEZ MARFIL	MARBELLA, 10,	Alcobendas	Madrid	C.A.Madri
11	982009100010627	KUKI	GATO	EUROPEO DE P	ADRIANA	VIDAL GARCIA-VALCARCEL	CERRO DE VALI	Las Rozas	Madrid	C.A.Madri
12	981100000481628	BRUNO	PERRO	PASTOR ALEM/	LUZ	DE HARO	CarreteraFUEN	Pozuelo	Madrid	C.A.Madri
13	981098102850133	RITA	PERRO	PASTOR ALEM/	ALFONSO	MORITZ	PEÑA PINTAD/	Madrid	Madrid	C.A.Madri
14	981098102800078	PATO	PERRO	LABRADOR RET	Mª ANGELES	SANCHEZ GUERRA MARI	ALFREDO MAR	Pinto	Madrid	C.A.Madri
15	981098102800035	ROMEO	PERRO	AKITA INU	TERESA	CANO MEDINA	NURIA, 49 1ª F	Pozuelo	Madrid	C.A.Madri
16	981098102798492	VEVA	GATO	EUROPEO DE P	SAAD	MARINA CUELI	PEÑA VIEJA, 9	Majadahonda	Madrid	C.A.Madri
17	981098102795975	PETER	GATO	PERSA	ANA	GARCIA DURAN	NURIA, 28 2ª F	Madrid	Madrid	C.A.Madri
18	981098102792945	ATILA	GATO	EUROPEO DE P	ASUNCION	KHALED	COSTA BRAVA,	Toledo	Toledo	Castilla la
19	981098102790634	PESI	PERRO	MESTIZO	JAVIER	GALDON FERNANDEZ	COSTA BRAVA,	Toledo	Toledo	Castilla la
20	981098102741912	POUPETTE	PERRO	FOX TERRIER P	MARTINEZ	GUERRERO	C/NURIA53	Cuenca	Cuenca	Castilla la
21	985120030950611	PANDORA	PERRO	CANICHE ENA	ABEL	CABRERA MARTINEZ	CERRO DEL CA	Cuenca	Cuenca	Castilla la
22	985120030981218	BASI	PERRO	PASTOR ALEM/	FELIPE	GARCIA GARCIA	PEÑA AUSEBA,	Guadalajara	Guadalajara	Castilla la
23	985120031039378	KATIA	GATO	SIAMÉS	ESTHER	FERNANDEZ DE LA ALDEA	NURIA, 20 CHA	Ciudad Real	Ciudad Real	Castilla la
24	985120031210386	LIS	PERRO	MESTIZO	CARMEN	GUIBELALDE	MORALZARZAL	Pozuelo	Madrid	C.A.Madri

Las tablas se organizan visualmente en filas y columnas: cada fila está dedicada a un registro, cada columna corresponde a un campo. Así el modelo entidad-relación de la anterior base de datos que contiene sólo una tabla llamada “Animales” sería el de la figura:



que desde el punto de vista de Access a esta definición le corresponde un diseño físico como, donde para cada tabla se declaran los campos:

Animales	
*	
id	
chip	
nombre	
especie	
raza	
nombre_propietario	
apellidos_propietario	
domicilio	
poblacion	
provincia	
comunidad	
codigo_postal	
saldo	

La única clave que podemos identificar es **id**, de ahí la llave junto al atributo. Aunque pudiera parecer que **chip** tiene un valor único para cada animal, no es posible asegurar que todos dispongan de uno.

La definición anterior es muy precaria desde el punto de vista de diseño, la iremos mejorando a lo largo de estos apuntes.

6. SELECT ... FROM ...

El tipo principal de consultas que se realizan a una base de datos es **consulta de selección** que permite realizar una recuperación de información. La especificación de la consulta se hace con un comando denominado SELECT cuya forma básica tiene el siguiente formato:

```
SELECT <lista de atributos>  
FROM <nombre tabla>
```

De la tabla indicada en la cláusula FROM la consulta recupera todos los registros, pero de ellos sólo los atributos indicados en la cláusula SELECT. Para recuperar todos los campos se utiliza el carácter *. Los campos estarán separados por comas. Los cambios de línea y separadores extra no son procesados.

Ejemplo:

```
SELECT chip, nombre FROM Animales
```

Devolvería el siguiente resultado:

chip	nombre
982009100150632	JAK
982009100141452	MARTINA
982009100141414	BRUTO
982009100141308	NAUS
982009100138878	BARTOK
982009100137490	PATY
982009100120235	NUQUI
982009100116535	KLAUS
982009100114324	POPI
982009100039117	DORA
982009100010627	KUKI
981100000481628	BRUNO
981098102850133	RITA
981098102800078	PATO
981098102800035	ROMEO
981098102798492	VEVA
981098102795975	PETER
981098102792945	ATILA
981098102790634	DESI

Cuando se buscan registros concretos que cumplan ciertas condiciones es necesario añadir al patrón anterior la cláusula WHERE que especifica la condición que deben cumplir los registros para ser seleccionados.

```
SELECT <lista de atributos>
FROM <nombre tabla>
WHERE <condición>
```

Ejemplo:

```
SELECT chip, nombre FROM Animales WHERE poblacion="Madrid"
```

Devolvería el siguiente resultado:

chip	nombre
981098102850133	RITA
981098102795975	PETER
985120032949725	MONA
985120032951521	LAIKA
DEU098100127103	CATHY

La condición puede ser tan complicada como sea necesario incluyendo múltiples operaciones con múltiples atributos, a **condición de que el resultado sea booleano**.

Nótese que no es necesario que los atributos que intervine en WHERE estén en la lista de atributos recuperados.

7. Índices

El trabajo habitual del SGBD, como la anterior consulta, es rebuscar en las tablas hasta encontrar registros con características definidas. Ese trabajo se puede realizar por fuerza bruta, inspeccionando uno a uno todos los registros de las tablas implicadas, lo que consumiría una cantidad de tiempo proporcional al número de registros (N) contenidos en la/s tabla/s, o puede seguirse una estrategia de ordenación de los datos que nos permita llegar a los registros buscados más rápidamente.

Si los registros se encontraran ordenados por el atributo implicado en la condición y sólo vayamos a extraer un registro, podríamos seguir la siguiente estrategia denominada búsqueda binaria:

1. Elegimos la tabla completa como la subtabla de búsqueda.
2. Localizar el registro que se encuentra en mitad de la subtabla
3. Si el registro es el buscado: Terminar
4. Si el registro buscado es anterior al localizado:
 - Elegir mitad anterior al registro de la subtabla como la nueva subtabla
 - Volver a 2
5. Si no (el registro buscado es posterior al localizado):
 - Elegir mitad posterior al registro de la subtabla como la nueva subtabla
 - Volver a 2

Será necesario añadir algunos detalles, por ejemplo, para contemplar la contingencia de que el registro buscado no existe en la tabla.

En este caso el tiempo consumido será del orden de $\log(N)$.

La cuestión entonces es tener ordenados los registros según el o los criterios necesarios para la aplicación que se le va a dar a la base. Las estructuras que registran estas ordenaciones se llaman **índices** y su elección y diseño son críticos a la hora de maximizar las prestaciones.

8. Mejorar el diseño

Recuperemos el diseño de la tabla Animales

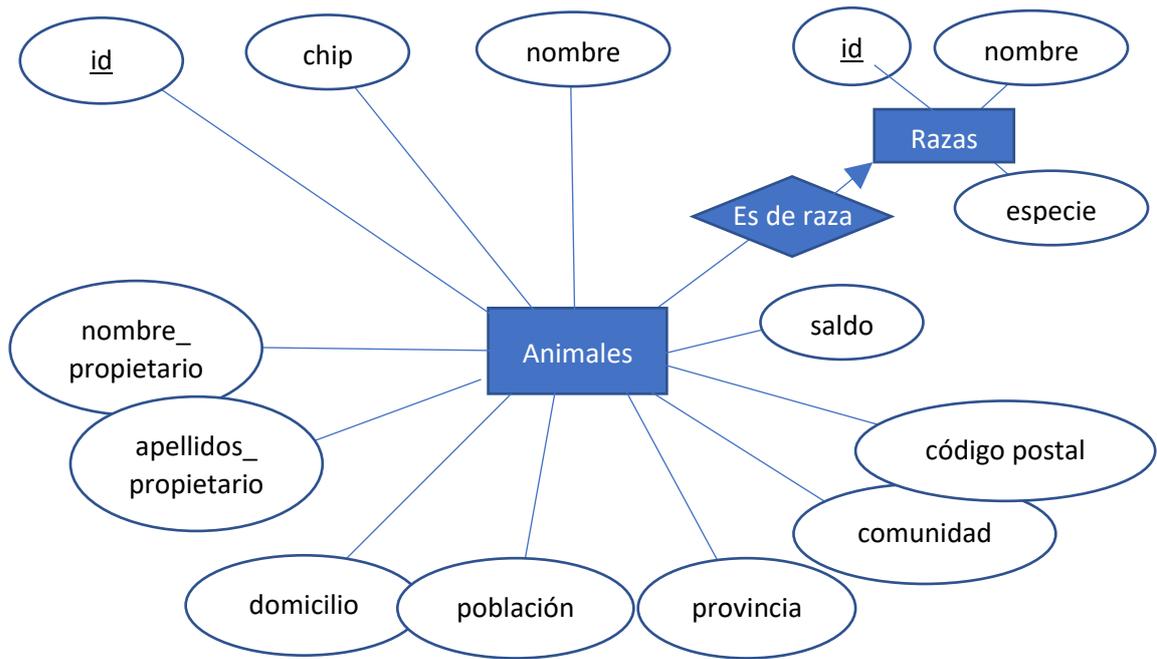
Animales	
*	
id	
chip	
nombre	
especie	
raza	
nombre_propietario	
apellidos_propietario	
domicilio	
poblacion	
provincia	
comunidad	
codigo_postal	
saldo	

y los datos

id	chip	nombre	especie	raza	nombre_propietario	apellidos_propietario	domicilio	poblacion	provincia	comunidad
1	982009100150632	JAK	PERRO	CHOW-CHOW	LEANDRO	LOPEZ VEIGUELA	AVDA. DE BETZ	Guadalajara	Guadalajara	Castilla la
2	982009100141452	MARTINA	GATO	SIAMÉS	CONCHA	GOMEZ CARDIN	AVDA. DE FRAI	Mostoles	Madrid	C.A.Madri
3	982009100141414	BRUTO	PERRO	MESTIZO	NOELIA	PERNAUTE	COLLADO DEL	Ciudad Real	Ciudad Real	Castilla la
4	982009100141308	NAUS	GATO	EUROPEO DE P	OLGA	PERNAUTE	VENTISQUERO	Alcorcón	Madrid	C.A.Madri
5	982009100138878	BARTOK	GATO	PERSA	NIEVES	RABASSA GARCIA	CARDENAL HEF	Cuenca	Cuenca	Castilla la
6	982009100137490	PATY	PERRO	COCKER SPANI	JOSE	CLAVIJO QUESADA	NURIA,16 CHAI	Toledo	Toledo	Castilla la
7	982009100120235	NUQUI	GATO	PERSA	PILAR	BOUZAS VITURRO	NURIA, 5 CHAL	Alcobendas	Madrid	C.A.Madri
8	982009100116535	KLAUS	PERRO	PASTOR ALEM/	DAVID	PALOMO RUIZ	NURIA, 91 ESC.	Las Rozas	Madrid	C.A.Madri
9	982009100114324	POPI	PERRO	PASTOR ALEM/	MERCEDES	RODRIGUEZ ANDIA-ROMER	ORTALEZA, 10€	Mostoles	Madrid	C.A.Madri
10	982009100039117	DORA	GATO	EUROPEO DE P	ALFONSO	GOMEZ MARFIL	MARBELLA, 10,	Alcobendas	Madrid	C.A.Madri
11	982009100010627	KUKI	GATO	EUROPEO DE P	ADRIANA	VIDAL GARCIA-VALCARCEL	CERRO DE VALI	Las Rozas	Madrid	C.A.Madri
12	981100000481628	BRUNO	PERRO	PASTOR ALEM/	LUZ	DE HARO	CarreteraFUEN	Pozuelo	Madrid	C.A.Madri
13	981098102850133	RITA	PERRO	PASTOR ALEM/	ALFONSO	MORITZ	PEÑA PINTAD/	Madrid	Madrid	C.A.Madri
14	981098102800078	PATO	PERRO	LABRADOR RET	Mª ANGELES	SANCHEZ GUERRA MARI	ALFREDO MAR	Pinto	Madrid	C.A.Madri
15	981098102800035	ROMEO	PERRO	AKITA INU	TERESA	CANO MEDINA	NURIA, 49 1º F	Pozuelo	Madrid	C.A.Madri
16	981098102798492	VEVA	GATO	EUROPEO DE P	SAAD	MARINA CUELI	PEÑA VIEJA, 9	Majadahonda	Madrid	C.A.Madri
17	981098102795975	PETER	GATO	PERSA	ANA	GARCIA DURAN	NURIA, 28 2ºF	Madrid	Madrid	C.A.Madri
18	981098102792945	ATILA	GATO	EUROPEO DE P	ASUNCION	KHALED	COSTA BRAVA,	Toledo	Toledo	Castilla la
19	981098102790634	PESI	PERRO	MESTIZO	JAVIER	GALDON FERNANDEZ	COSTA BRAVA,	Toledo	Toledo	Castilla la
20	981098102741912	POUPETTE	PERRO	FOX TERRIER P	MARTINEZ	GUERRERO	C/NURIA53	Cuenca	Cuenca	Castilla la
21	985120030950611	PANDORA	PERRO	CANICHE ENA	ABEL	CABRERA MARTINEZ	CERRO DEL CA	Cuenca	Cuenca	Castilla la
22	985120030981218	BASI	PERRO	PASTOR ALEM/	FELIPE	GARCIA GARCIA	PEÑA AUSEBA,	Guadalajara	Guadalajara	Castilla la
23	985120031039378	KATIA	GATO	SIAMÉS	ESTHER	FERNANDEZ DE LA ALDEA	NURIA, 20 CHA	Ciudad Real	Ciudad Real	Castilla la
24	985120031210386	LIS	PERRO	MESTIZO	CARMEN	GUIBELALDE	MORALZARZAL	Pozuelo	Madrid	C.A.Madri

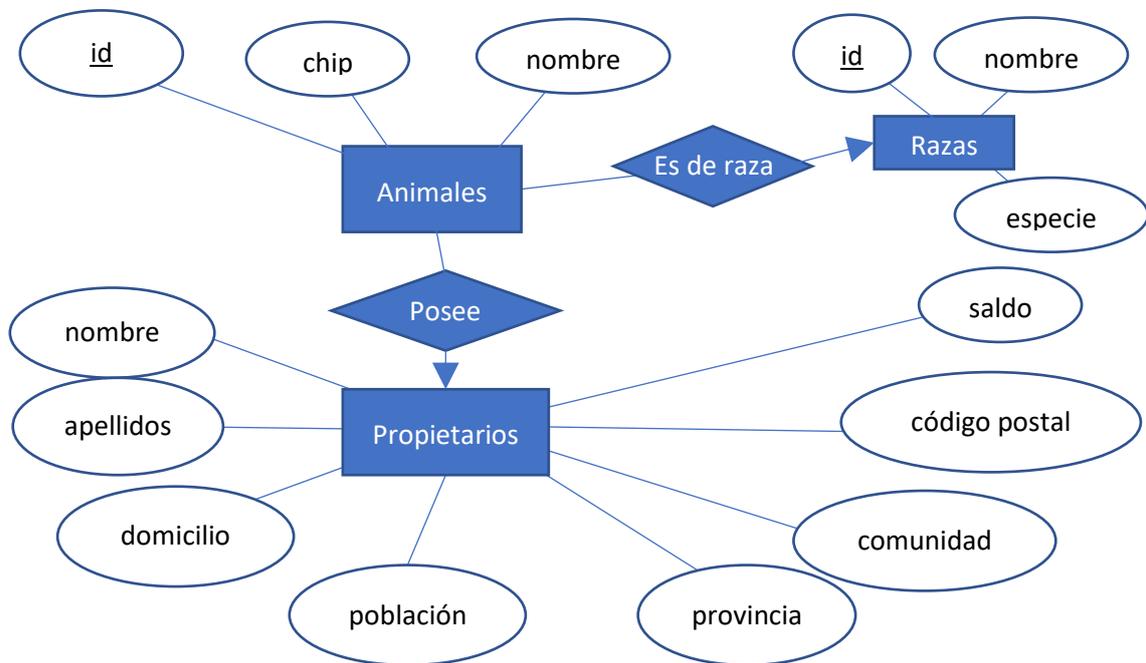
Si nos fijamos en los registros 8, 9 y 22 vemos que los tres animales son de raza “Pastor Alemán” y en los tres casos la especie es “Perro”. No puede ser de otra forma porque es obligado que todos los “Pastores Alemanes” sean perros, lo mismo ocurre con cualquier otra raza de perros o de gatos.

Se trata de una redundancia explícita: teniendo la raza, la especie no añade ninguna información, pero sí es necesario conocer esa relación. Con esa idea vamos a cambiar el diseño conceptual (Modelo E-R)

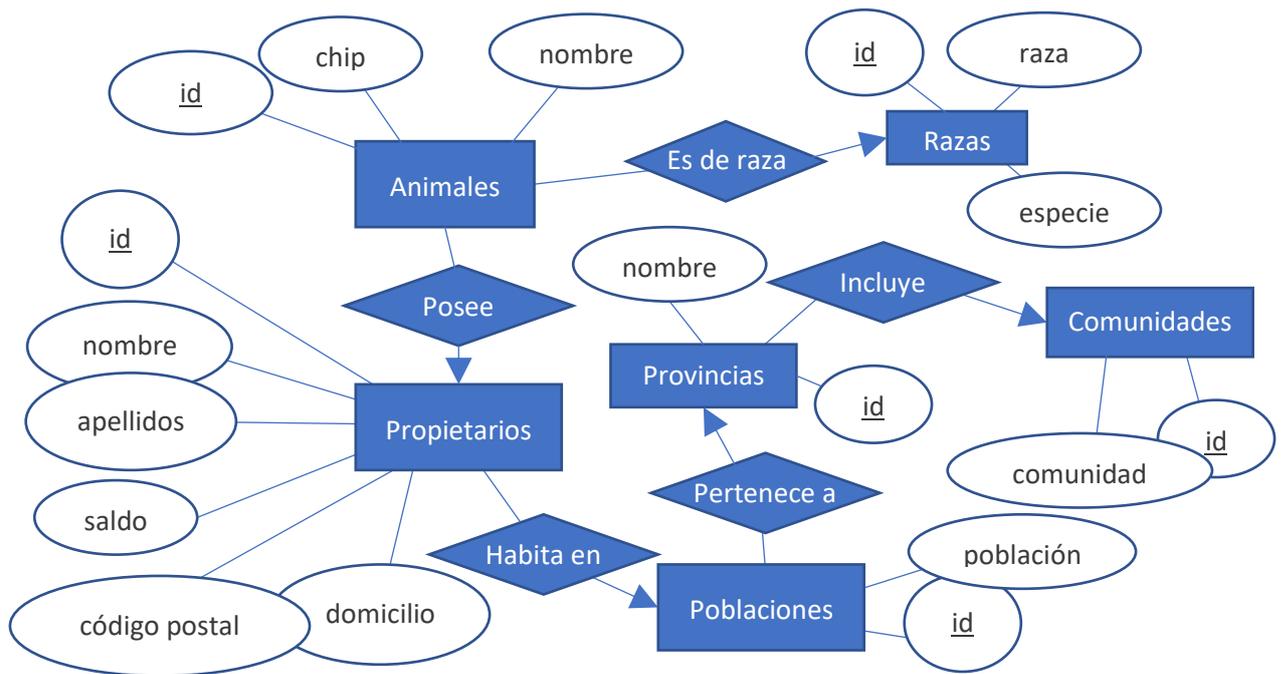


No podemos asegurar que el nombre de la raza sea clave, ya que otras especies pueden tener razas con ese nombre, por tanto, se añade un **id** como clave artificial.

Lo mismo ocurre con los datos de domicilio y saldo, cualquier otra mascota del mismo dueño tendrá los mismos valores en estos atributos: redundancia



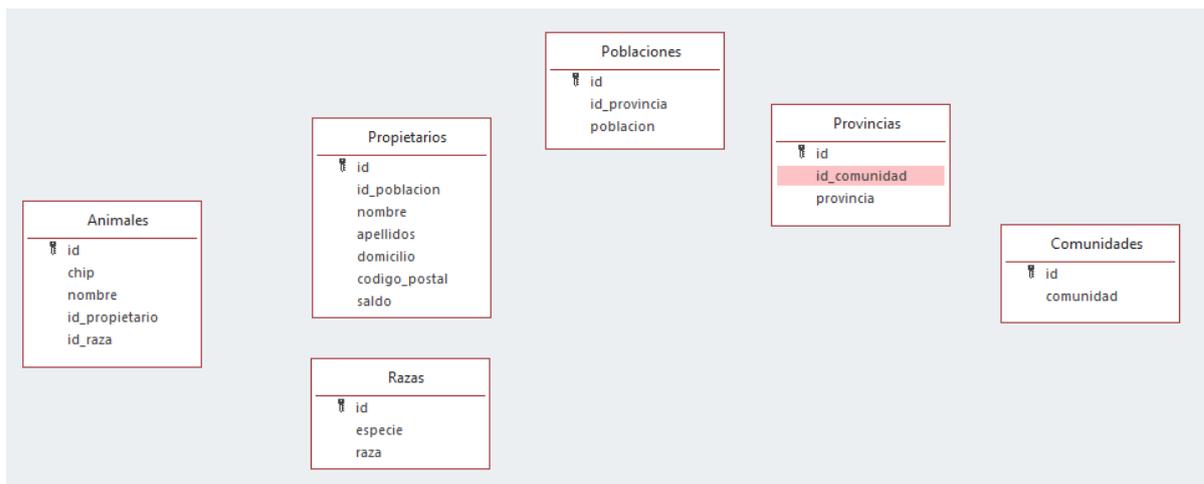
Lo mismo ocurre con población, provincia y comunidad.



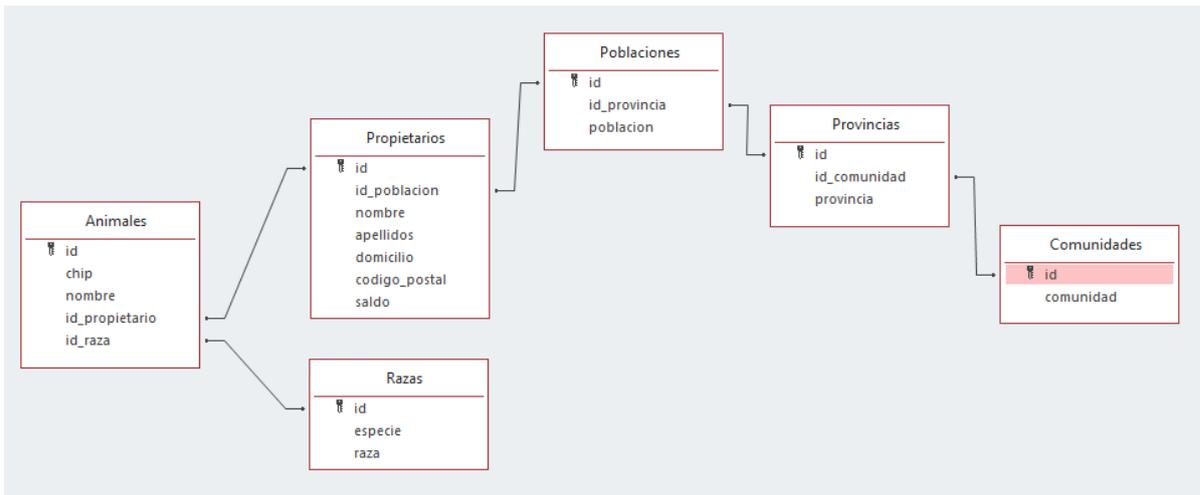
9. Relaciones

La forma de implementar las relaciones de cardinalidad 1 a n es añadir un campo a la tabla de la entidad n con el valor de la clave de la entidad del 1. Así en Animales habría que añadir el id del propietario y el id de la raza. Esas claves establecen el vínculo que relacionan las tablas. Lo mismo con poblaciones y el id de su provincia, provincias y el id de sus comunidades, propietarios y el id de sus poblaciones.

El diseño físico empezaría por determinar las tablas y sus campos. Sistemáticamente en todas las tablas se incluye un campo **id** que sirva de clave (valores únicos)



Y añadiendo los vínculos que representan las relaciones



Donde las relaciones se establecen vinculando los atributos.

Animales.id_propietario → Propietarios.id (n a 1)

Animales.id_raza → Razas.id (n a 1)

Propietarios.id_poblacion → Poblaciones.id (n a 1)

Poblaciones.id_provincia → Provincias.id (n a 1)

Provincias.id_comunidad → Comunidades.id (n a 1)

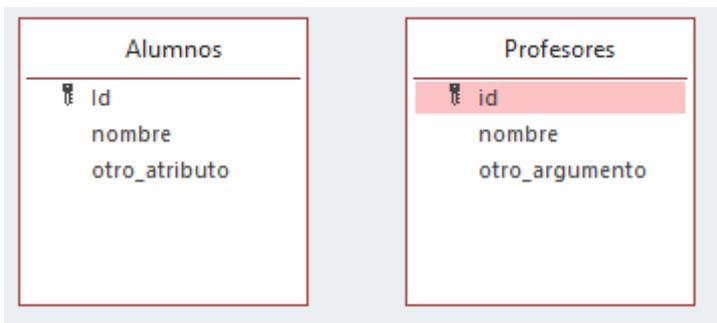
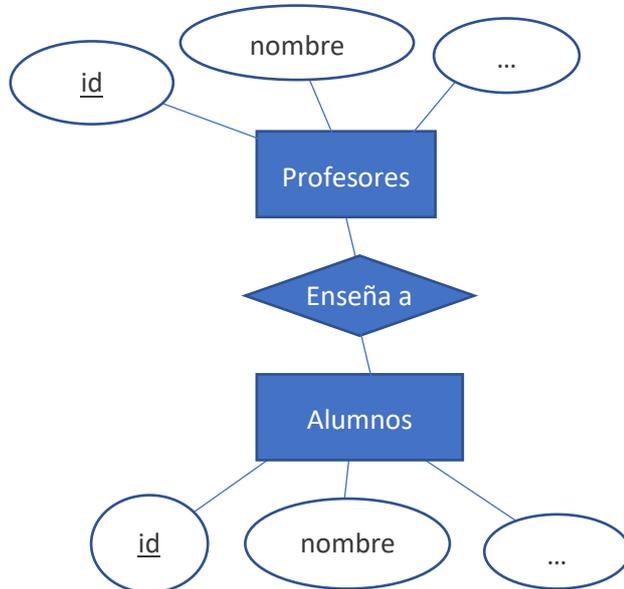
Se reconoce que la relación es n a 1 porque en el lado del 1 utiliza para el vínculo un atributo clave (cuyo valor es único) mientras que en el lado del n no.

10. Relaciones 1 a 1

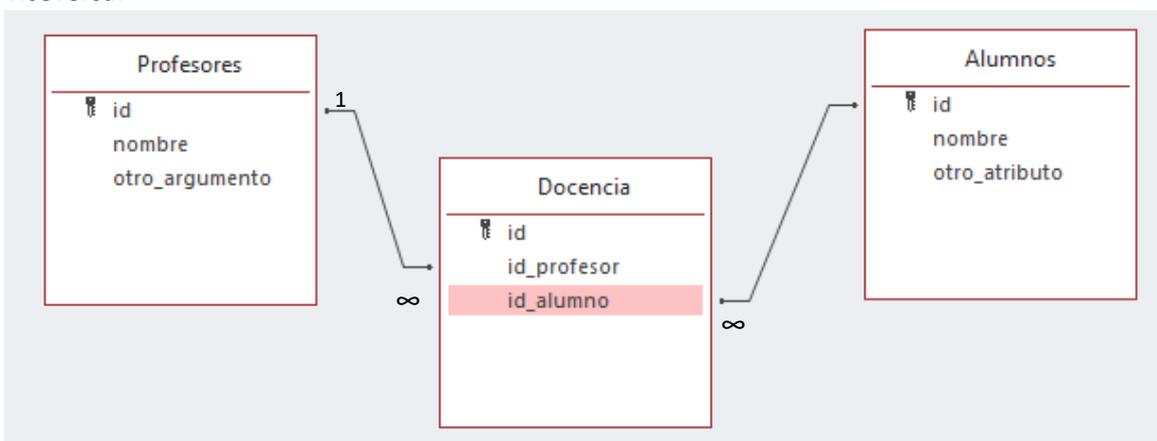
Si fuera necesario crear una relación 1 a 1 el procedimiento es el mismo que para las relaciones 1 a n, asumiendo que uno de los dos 1 es n. Será necesario añadir a una de las tablas un nuevo campo para vincular con el id de la tabla relacionada.

11. Relaciones n a n

Este es un problema distinto, ya que con vínculos sólo podemos relacionar un registro de cada tabla.



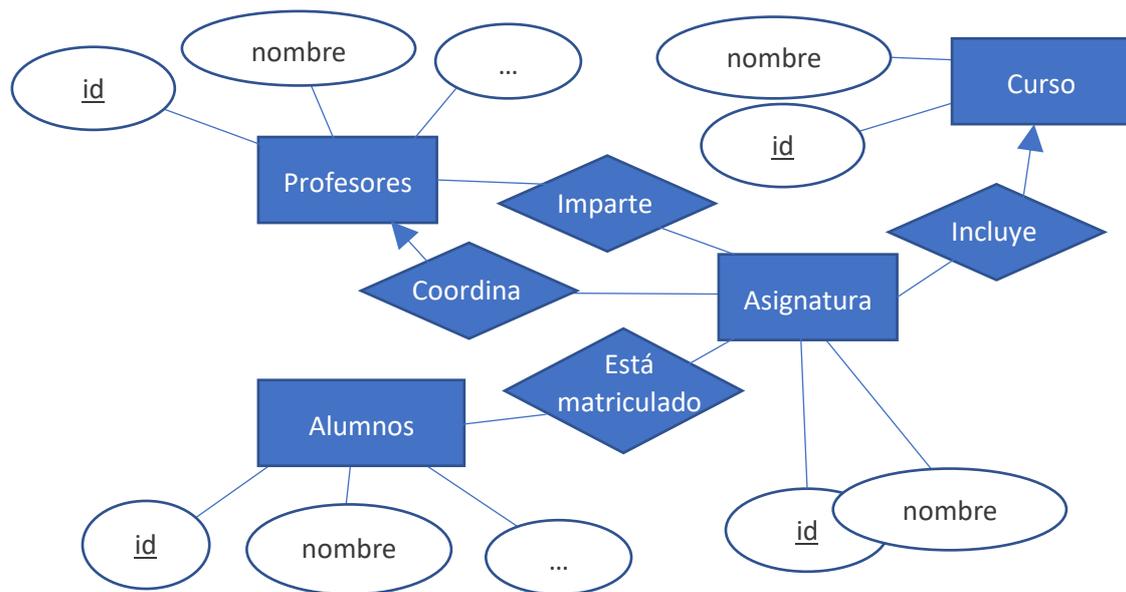
La solución es crear una tabla que permita relacionar a cualquier alumno con cualquier profesor y viceversa.



La relación Docencias-Profesores es 1 a n. La relación Docencias-Alumnos es 1 a n

Leído de forma inversa un profesor tiene múltiples docencias cada una con un alumno y por tanto un profesor tiene múltiples docencias con múltiples alumnos. En sentido contrario un alumno está relacionado con múltiples docencias cada una con un profesor, en consecuencia, cada alumno está relacionado docentemente con múltiples profesores, con lo que queda implementada la relación Profesor – Alumno (n a n).

12. Ejercicios



Realizar el diseño físico de la base de datos a partir del diseño conceptual anterior.