

Introducción a MongoDB

Pablo Hidalgo

Contents

1	Instalación	1
2	Primeros pasos	1
3	Leyendo datos	3
3.1	Jerarquías de datos	4
3.2	Operaciones OR	4
3.3	Actualizando datos	5
3.4	Eliminar datos	6
3.5	Eliminar una colección y una base de datos	6

MongoDB es una de las bases de datos NoSQL más utilizada. Ya hemos visto en clase que se trata de una base de datos orientada a documentos.

1 Instalación

La instalación para los sistemas operativos soportados está documentada aquí.

Al final, pregunta se quiere instalar **MongoDB Compass**. Se trata de una interfaz gráfica que permite explorar los datos de forma visual y hacer algunas operaciones sencillas. Sin embargo, su capacidad es limitada. En este curso nos centraremos en el uso de la ventana de comandos

2 Primeros pasos

Una vez que se ha completado la instalación, ya podemos empezar a *divertirnos* con MongoDB.

Si estás en Windows, busca el ejecutable **mongo.exe** (Lo habitual es que lo puedas encontrar en **C:\Program Files\MongoDB\Server\4.0\bin\mongo.exe**).

Si estás en MacOS,

Asegúrate de que MongoDB está funcionando. Para ello, debes haber ejecutado en la línea de comandos **mongod** y tienes que haber obtenido algo como **initandlisten] waiting for connections on port 27017**. Ten cuidado de no cerrar esta ventana ya que matarías la sesión de MongoDB.

Abre una nueva ventana de la línea de comandos y escribe

```
mongo master
```

Esta sentencia genera una conexión al *motor* de MongoDB y crea una base de datos **master** en el caso de que no exista previamente.

Escribe **help** para ver qué sentencias puedes utilizar.

Como nos hemos conectado a MongoDB con la sentencia **mongo master** estamos en la base de datos **master**. Podemos ver qué otras bases de datos tenemos disponibles escribiendo **show dbs** y cambiar a otra base de datos escribiendo **use nombre_bbdd**.

Ahora mismo nuestra base de datos está vacía; tenemos que empezar a rellenarla con datos. Para crear una colección (recuerda la nomenclatura en Mongo y piensa en que las colecciones son análogas a las tablas en una base de datos relacional) es tan fácil como empezar a añadir registros, ya que no es necesario crear una estructura (ni un esquema) simplemente hay que empezar a usarla. Además, si te diste cuenta antes, cuando ejecutamos `show dbs` no aparecía por ningún lado nuestra base de datos `master` ya que ésta no existe propiamente hasta que no comienza a tener valores. Para crear una colección `compras` tenemos que **insertar** un valor con ayuda del comando `db.compras.insert()`:

```
db.compras.insert({
  nombre: "Pablo Hidalgo",
  dni: 12345,
  fecha_nacimiento: ISODate("1991-02-09"),
  edad: 28,
  articulos: [150, 756, 312],
  direccion: {
    calle: "Príncipe de Vergara",
    numero: 43,
    ciudad: "Madrid"
  }
})
```

Como hemos visto en clase, los documentos (recuerda, una colección está compuesta de documentos, algo similar a una fila o un registro en una base de datos relacional) siguen un formato al estilo JSON (estrictamente, se tratan de BSON) por lo que cualquier nuevo documento que añadamos debe seguir ese formato donde las llaves `{...}` denotan un objeto compuesto por parejas de atributo-valor y donde `[...]` significa un array.

Si te fijas en el documento que acabamos de insertar, uno de los valores es `articulos` al que lo hemos declarado como un *array* cosa que, por construcción, no era posible hacer en una base de datos relacional ya que uno de los principios es que una celda (la intersección entre una fila y una columna) solo podía contener un único valor. Eso nos podía llevar en algún caso a tener alguna información duplicada. Además, una propiedad relevante de los documentos es que se introduce de forma natural el concepto de jerarquía como puedes ver en el atributo `direccion`.

Ahora podemos verificar que existe nuestra base de datos con `show dbs` y ver la colección con `show collections`. Podemos listar el contenido de una colección con `find()`:

```
db.compras.find()
```

Como puedes ver, la sintaxis es muy diferente a la misma operación escrita en lenguaje SQL:

```
SELECT *
FROM compras;
```

Como habrás observado, el documento contiene un campo `_id` que Mongo genera automáticamente y que se trata de un identificador único, algo así como el datatype `SERIAL` en PostgreSQL. Este `_id` siempre tiene 12 bytes y está compuesto por un *timestamp*, un identificador de la máquina cliente, un identificador del proceso del cliente y un contador de 3-bytes incremental.

Aunque pueda parecer que para nosotros, como usuarios, cómo se genere `_id` no es relevante, lo cierto es que esa forma de generarlo permite que cada máquina no tenga que conocer los identificadores que hayan generado otras siendo fundamental para poder lanzar procesos distribuidos (el *datatype* `SERIAL` de PostgreSQL necesita conocer el último generado para incrementarlo en una unidad).

Quando ejecutamos la función `find()`, la consola muestra cada documento en una sola línea que hace difícil la visualización. Puedes añadir al final `.pretty` para obtener una vista más intuitiva. Pruébalo con `db.ciudades.find().pretty()`

Por completar nuestra colección, vamos a añadir algún documento más. Para añadir varios documentos a la vez en una colección, utilizamos `[...]`:

```

db.compras.insert(
[
  {
    nombre: "Andrea Pérez",
    dni: 54321,
    fecha_nacimiento: ISODate("1985-03-12"),
    edad: 34,
    articulos: [715, 1],
    direccion: {
      calle: "Alcalá",
      numero: 69,
      ciudad: "Madrid"
    }
  },
  {
    nombre: "Juan García",
    dni: 11111,
    fecha_nacimiento: ISODate("1975-10-01"),
    edad: 44,
    articulos: [635, 4441],
    direccion: {
      calle: "Cervantes",
      numero: 3,
      ciudad: "Alcorcón"
    }
  },
  {
    nombre: "Nerea Álvarez",
    dni: 22222,
    fecha_nacimiento: ISODate("1995-06-29"),
    edad: 24,
    direccion: {
      calle: "Nueva",
      numero: 1,
      ciudad: "Alcobendas"
    }
  }
]
)

```

Una función que puede resultar útil es saber cuántos documentos existen dentro de una colección. Esto lo podemos hacer con la función `count()`:

```
db.compras.count()
```

3 Leyendo datos

Ya hemos visto que la función `find()` sin parámetros nos da todos los documentos. Para especificar unos criterios de búsqueda, tenemos que hacerlo de la siguiente forma

```
db.compras.find({dni: 12345})
```

La función `find()` acepta un segundo parámetro opcional con el que especificar qué campos queremos obtener. Por ejemplo, si solo queremos el nombre (además del `_id`), podemos escribir:

```
db.compras.find({dni: 12345}, {nombre: true})
```

Si queremos obtener todos los campos excepto el nombre, podemos hacerlo con:

```
db.compras.find({dni: 12345}, {nombre: false})
```

¿Cómo harías para devolver los campos `nombre` y `fecha_nacimiento` del registro con DNI 12345?

```
db.compras.find({dni: 12345}, {nombre: true, fecha_nacimiento: true})
```

Igual que en PostgreSQL, en Mongo podemos construir queries específicas que nos permitan encontrar aquellos documentos que cumplan algún requisito.

Por ejemplo, para encontrar aquellos clientes mayores de 30 años debemos escribir:

```
db.compras.find({edad: {$gt: 30}})
```

En Mongo, los operadores condicionales se escribe como `{$operador: valor}`.

Aquí se pueden consultar los operadores que existen en Mongo. Algunos habituales son:

- `$eq`: Valores **idénticos** a un valor especificado.
- `$gt`: Valores **mayores** que un valor especificado.
- `$lt`: Valores **menores** que un valor especificado.
- `$gte`: Valores **mayores o iguales** que un valor especificado.
- `$lte`: Valores **menores o iguales** que un valor especificado.
- `$in`: valores en un **array** de valores.
- `$ne`: Valores **distintos** a un valor especificado.

Encuentra aquellos clientes cuya edad sea menor o igual que 34 años.

3.1 Jerarquías de datos

Mongo está muy bien preparado para trabajar con jerarquías de datos. Para filtrar *subdocumentos*, hay que separar los campos anidados por puntos. Por ejemplo, para encontrar aquellos clientes que viven en Madrid, lo haríamos

```
db.compras.find(
  {"direccion.ciudad" : "Madrid"}
)
```

3.2 Operaciones OR

Cuando utilizamos la función `find()` estamos utilizando implícitamente operaciones **and**, es decir, solo devuelve aquellos documentos que cumplan **todos** los criterios que hayamos especificado. Por ejemplo, la siguiente sentencia no devuelve ningún resultado ya que ningún documento satisface las dos condiciones **a la vez**:

```
db.compras.find(
  {
    "edad": 24,
    "direccion.ciudad": "Madrid"
  }
)
```

Sin embargo, si queremos obtener aquellos registros que satisfacen *alguna* de las condiciones, necesitamos utilizar el operador **or**:

```
db.compras.find(
  {
    $or: [
      {"edad": 24},
      {"direccion.ciudad": "Madrid"}
    ]
  }
)
```

Ejercicio: encuentra aquellos clientes que vivan en Madrid o en Alcobendas. ¿Sabrías hacerlo de dos formas distintas?

```
db.compras.find(
  {
    $or: [
      {"direccion.ciudad": "Madrid"},
      {"direccion.ciudad": "Alcobendas"}
    ]
  }
)
```

```
db.compras.find({"direccion.ciudad": {$in: ["Madrid", "Alcobendas"]}})
```

3.3 Actualizando datos

Hasta ahora hemos añadido nuevos documentos a una colección. Otra tarea que queremos saber hacer es **cómo actualizar documentos ya existentes**. Para ello utilizamos la función `update(criterio, operacion)`. El primero de los parámetros, `criterio`, es el mismo tipo de objeto que se le pasa a la función `find()`; el segundo parámetro, `operacion`, es un objeto cuyos campos reemplazarán a aquellos documentos que concuerden con lo escrito en `criterio`.

Por ejemplo, si quisiésemos actualizar el DNI y la fecha de nacimiento de un cliente, tendríamos que

```
db.compras.update(
  {nombre: "Pablo Hidalgo"},
  {$set : {dni: 11345, fecha_nacimiento: ISODate("1991-02-10")}}
)
```

Quizá te preguntes si es necesaria escribir `$set`. Si en vez de la sentencia anterior escribiésemos

```
db.compras.update(
  {nombre: "Pablo Hidalgo"},
  {dni: 11345, fecha_nacimiento: ISODate("1991-02-10")}
)
```

Mongo entendería que lo que quieres es **reemplazar por completo** el documento. Así que, presta atención cuando quieras actualizar datos.

```
Si has ejecutado la última sentencia, puedes volver al estado original ejecutando db.compras.update(
  {dni: 11345}, { nombre: "Pablo Hidalgo", dni: 12345, fecha_nacimiento: ISODate("1991-02-09"),
  edad: 28, articulos: [150, 756, 312], direccion: { calle: "Príncipe de Vergara",
  numero: 43, ciudad: "Madrid" } })
```

Podemos hacer más cosas que modificar un determinado valor. Por ejemplo, `$inc` nos permite incrementar un número por una cantidad determinada:

```
db.compras.update(
  {dni: 11345},
```

```
{ $inc: { edad: 1 } }  
)
```

Se van añadiendo constantemente nuevos operadores, por lo que es aconsejable consultar la documentación para estar al día de los cambios. Algunos relevantes son:

- **\$set**: cambia un campo con un valor dado.
- **\$unset**: elimina un campo.
- **\$inc**: incrementa un campo por un número dado.
- **\$pop**: elimina el último (o el primer) elemento de un array.
- **\$push**: añade un elemento a un array.

Ejercicio: El cliente con DNI 11111 ha comprado el artículo 1. Actualiza la base de datos para reflejar esta compra.

```
db.compras.update(  
  {dni: 11111},  
  {  
    $push: {articulos: 1}  
  }  
)
```

3.4 Eliminar datos

Otra operación habitual es la de **eliminar documentos** de una colección. En Mongo esto se puede hacer simplemente reemplazando **find()** por **remove()** y todos los documentos que satisfagan la condición se eliminarán.

Antes de eliminar cualquier documento, es recomendable ejecutar **find()** para comprobar que el criterio está bien construido y hace lo que queremos que haga.

Ejercicio: Elimina todos los clientes que vivan en Madrid.

3.5 Eliminar una colección y una base de datos

Por último, dos tareas que pueden ser útiles son las de eliminar una colección entera (junto con todos sus documentos) y eliminar completamente una base de datos.

Para eliminar una colección utilizamos **drop()**. Para comprobar su funcionamiento, primero vamos a asegurarnos de las colecciones existentes en nuestra base de datos:

```
show collections
```

Para eliminar la colección **compras** podemos ejecutar

```
db.ciudades.drop()
```

Para eliminar una base de datos completa, necesitamos utilizar **dropDatabase()**. Igual que en el caso anterior, primero mostramos las bases de datos disponibles:

```
show dbs
```

y no situamos en aquella que queramos borrar (por ejemplo, **use master**).

Y, para borrar la base de datos que estemos usando, escribimos:

```
db.dropDatabase()
```