

UNIVERSIDAD DE CUENCA

# Consultas en MongoDB - II

BASE DE DATOS II

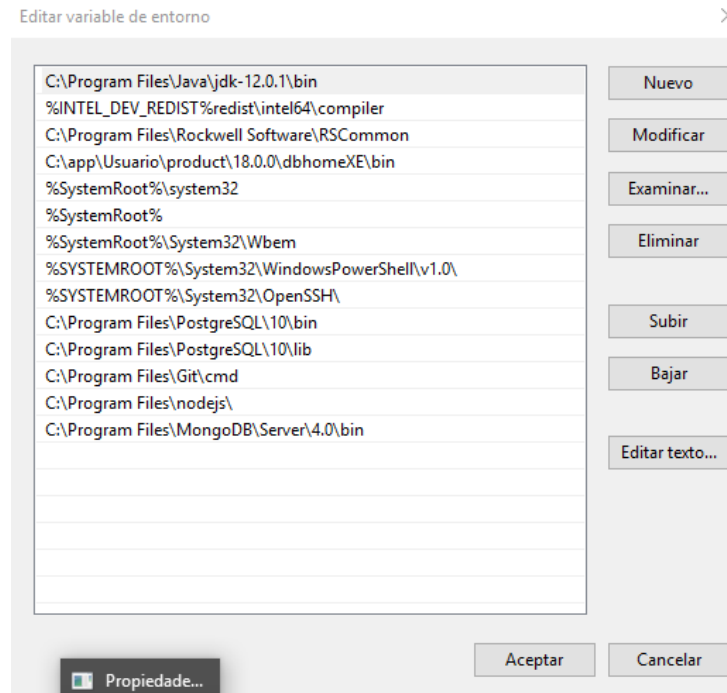
Freddy Abad & David Santos  
26 de junio de 2019

## PRACTICA 2

La instalación de Mongo DB para el servidor y para el cliente, tiene el proceso normal de cualquier programa .exe o .msi. El tutorial donde nos basamos para realizar la instalación de este SGBD, fue tomado de un video en YouTube:

*How to Install MongoDB on Windows 10 – ProgrammingKnowledge.* Disponible en: <https://www.youtube.com/watch?v=FwMwO8pXfq0>

Para facilitar el trabajo en este SGBD, se modifica el PATH



Para la creación de una Base de Datos llamada practicaBD2 se procede con los comandos:

```
use practicaBD2
show collections
```

Además, crear una tabla/collection llamada practicaJulio, la base de datos tiene más de 1250000 registros ingresados en formato JSON. El tutorial en el cual nos basamos para esta parte de la práctica es el siguiente:

*Import JSON data in mongodb collection – Satyendra\_kumar.* Disponible en: <https://www.youtube.com/watch?v=Sl-UFrpDYKU>

El comando a utilizar se define a continuación:

```
mongoimport --db practicaBD2 --collection practicaJulio --file
C:\Users\Usuario\Desktop\mongodataset.json
```

```
C:\Program Files\MongoDB\Server\4.0\bin>mongoimport --db practicaBD2 --collection practicaJulio --file C:\Users\Usuario\Desktop\mongodataset.json
2019-07-03T11:32:30.133-0500 connected to: localhost
2019-07-03T11:32:32.100-0500 [#.....] practicaBD2.practicaJulio 23.8MB/523MB (4.5%)
2019-07-03T11:32:35.100-0500 [###.....] practicaBD2.practicaJulio 71.2MB/523MB (13.6%)
2019-07-03T11:32:38.100-0500 [####.....] practicaBD2.practicaJulio 119MB/523MB (22.7%)
2019-07-03T11:32:41.100-0500 [#####.....] practicaBD2.practicaJulio 166MB/523MB (31.8%)
2019-07-03T11:32:44.100-0500 [#####.....] practicaBD2.practicaJulio 211MB/523MB (40.4%)
2019-07-03T11:32:47.100-0500 [#####.....] practicaBD2.practicaJulio 253MB/523MB (48.5%)
2019-07-03T11:32:50.101-0500 [#####.....] practicaBD2.practicaJulio 293MB/523MB (56.1%)
2019-07-03T11:32:53.100-0500 [#####.....] practicaBD2.practicaJulio 334MB/523MB (64.0%)
2019-07-03T11:32:56.100-0500 [#####.....] practicaBD2.practicaJulio 381MB/523MB (72.8%)
2019-07-03T11:32:59.100-0500 [#####.....] practicaBD2.practicaJulio 431MB/523MB (82.4%)
2019-07-03T11:33:02.101-0500 [#####.....] practicaBD2.practicaJulio 474MB/523MB (90.7%)
2019-07-03T11:33:05.100-0500 [#####.....] practicaBD2.practicaJulio 501MB/523MB (95.7%)
2019-07-03T11:33:07.056-0500 [#####.....] practicaBD2.practicaJulio 523MB/523MB (100.0%)
2019-07-03T11:33:07.056-0500 imported 418932 documents
```

Número de registros ingresados.

```
type "it" for more
> db.practicaJulio.count()
1256796
```

Peso de la base de datos:

```
> show dbs
admin          0.000GB
config         0.000GB
local          0.000GB
mylibs         0.000GB
practicaBD2    0.836GB
test           0.000GB
```

Descripción General Base de Datos:

La estructura de la BD se muestra a continuación, se utiliza el comando `.pretty()` para visualizar de manera amigable:

```
> db.practicaJulio.find().pretty();
{
  "_id" : ObjectId("5d1cd5d5f007640a1f21b510"),
  "title" : "Specification by Example",
  "isbn" : "1617290084",
  "pageCount" : 0,
  "publishedDate" : ISODate("2011-06-03T07:00:00Z"),
  "thumbnailUrl" : "https://s3.amazonaws.com/AKIAJC5RLADLUMVRPFDQ.book-thumb-images/adzic.jpg",
  "status" : "PUBLISH",
  "authors" : [
    "Gojko Adzic"
  ],
  "categories" : [
    "Software Engineering"
  ]
}
```

Para la realización de las consultas en la Base de Datos, no guiamos en la documentación oficial del SGBD: *MongoDB Documentation*. Disponible en:

<https://docs.mongodb.com/manual/reference/method/db.collection.find/>

SELECT simple con una condición, mostrar los registros con numero de páginas igual a 600:

```
db.practicaJulio.find({"pageCount" : 600}).count();
> db.practicaJulio.find({"pageCount" : 600}).count();
17496
```

SELECT simple con varias condiciones, muestra el número de los registros con número de páginas igual a 600 y categorías Comerciales:

```
db.practicaJulio.find({"pageCount" : 600}, {"categories": "Business"}).pretty();
> db.practicaJulio.find({"pageCount" : 600, "categories": ["Internet"]}).count();
2916
```

SELECT simple con varias condiciones, muestra el número total de los registros con numero de paginas igual a 600 y fecha de publicación igual a 1/Julio/2007:

```
db.practicaJulio.find({"pageCount" : 600, "publishedDate" : ISODate("2007-07-01T07:00:00Z")}).count();
> db.practicaJulio.find({"pageCount" : 600, "publishedDate" : ISODate("2007-07-01T07:00:00Z")}).count();
2916
```

SELECT simple con varias condiciones, mostrar los registros con numero de paginas igual a 600 y fecha de publicación igual a 1/Julio/2007:

```
db.practicaJulio.find({"pageCount" : 600, "publishedDate" : ISODate("2007-07-01T07:00:00Z")}).pretty();
{
  "_id" : ObjectId("5d1cd5d6f007640a1f21d248"),
  "title" : "Ant in Action",
  "isbn" : "193239480X",
  "pageCount" : 600,
  "publishedDate" : ISODate("2007-07-01T07:00:00Z"),
  "thumbnailUrl" : "https://s3.amazonaws.com/AKIAJCSRLADLUMVRPFDQ.book-thumb-images/loughran.jpg",
  "shortDescription" : "The most widely used build tool for Java projects, Ant is cross-platform, extensible, simple, and fast. It scales from small personal projects to large, multi-team enterprise projects. And, most important, it's easy to learn.",
  "longDescription" : "Ant in Action is a complete guide to using Ant to build, test, redistribute and deploy Java applications. A retitled second edition of the bestselling and award-winning Java Development with Ant, this book contains over 50% new content including: New Ant 1.7 features Scalable builds for big projects Continuous integration techniques Deployment Library management Extending Ant Whether you are dealing with a small library or a complex server-side system, this book will help you master your build process. By presenting a running example that grows in complexity, the book covers nearly the entire gamut of modern Java application development, including test-driven development and even how to set up your database as part of the deployment.",
  "status" : "PUBLISH",
  "authors" : [
    "Steve Loughran",
    "Erik Hatcher"
  ],
}
```

### Tiempos de consulta

Existen dos posibilidades de revisar el rendimiento temporal de la base de datos, la primera nos da valores promedio generales que toman, estos datos son almacenados en una tabla del sistema:

Comando a utilizar:

```
db.system.profile.find({"millis":{$gt:10}}, {"millis:NumberInt(1), "query":NumberInt(1), "command.pipeline":1}).sort({"millis":-1})
```

Resultado:

```
{ "command" : { }, "millis" : 100, "query" : 1, "command.pipeline" : 1 }
> db.system.profile.find({"millis":{$gt:10}}, {"millis:NumberInt(1), "query":NumberInt(1), "command.pipeline":1}).sort({"millis":-1})
{ "command" : { }, "millis" : 224 }
{ "command" : { }, "millis" : 32 }
{ "command" : { }, "millis" : 30 }
```

Estos datos nos indica valor máximo, valor mínimo y valor promedio de todas las consultas que se han realizado, sucesivamente.

La segunda opción a utilizar es el comando:

```
db.setProfilingLevel(2);
```

Este comando se debe ejecutar luego de cada consulta realizada. Como resultado obtendremos, caso ejemplo, donde slowms, nos da el tiempo en milisegundos que toma la consulta:

```
> db.setProfilingLevel(2);
{ "was" : 2, "slowms" : 100, "sampleRate" : 1, "ok" : 1 }
```

Querys a trackear:

```
db.practicaJulio.find({"pageCount" : 600}).pretty();
db.practicaJulio.find({"pageCount" : 600,"categories":"Internet"}).pretty();
db.practicaJulio.find({"pageCount" : 600,"categories":["Internet"]}).count();
db.practicaJulio.find({"pageCount" : 600,"publishedDate" : ISODate("2007-07-01T07:00:00Z")}).count();
db.practicaJulio.find({"pageCount" : 600,"publishedDate" : ISODate("2007-07-01T07:00:00Z")}).pretty();
```

Resultado

Query	Tiempo (ms)
db.practicaJulio.find({"pageCount" : 600}).pretty();	100
db.practicaJulio.find({"pageCount" : 600,"categories":"Internet"}).pretty();	110
db.practicaJulio.find({"pageCount" : 600,"publishedDate" : ISODate("2007-07-01T07:00:00Z")}).count();	105
db.practicaJulio.find({"pageCount" : 600,"publishedDate" : ISODate("2007-07-01T07:00:00Z")}).count();	102
db.practicaJulio.find({"pageCount" : 600,"publishedDate" : ISODate("2007-07-01T07:00:00Z")}).pretty();	103