



#### ACTIVIDADES POR DESARROLLAR:

1. Introducción a las bases de datos no relacionales (NOSQL)
2. instalación de un sistema manejador de base de datos no relacional (MONGODB)
3. Comandos utilizados en MONGODB
4. Migrar base de datos relacionales a no relacional
5. Implementar desde PYTHON la conectividad hacia MONGODB

#### EVIDENCIA(S) A ENTREGAR:

1. Escribe un ensayo sobre que son las bases de datos NOSQL y aplicaciones existentes.
2. Elabora informe de la ejecución del plan de instalación de MONGODB.
3. Realiza los ejercicios propuestos en el presente taller.
4. Realiza la conversión de la base de datos SALUD hacia MONGODB y entrega el script de creación y el respectivo informe.
5. Desarrolla en Python un programa que se conecte a una base de datos MONGODB y a una colección mostrando por pantalla los datos guardados.

**DURACIÓN: 24 HORAS 16 horas trabajo autónomo.**

#### CONTROL DEL DOCUMENTO

	Nombre	Cargo	Dependencia	Fecha
Autor (es)	JOSE FERNANDO GALINDO SUAREZ	INSTRUCTOR	CGMLTI	14/09/2022

**CONTROL DE CAMBIOS** (diligenciar únicamente si realizan ajustes al taller)

	Nombre	Cargo	Dependencia	Fecha	Razón del Cambio
Autor (es)					

## PRESENTACIÓN

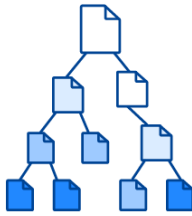
El auge de internet ha propiciado la generación de grandes volúmenes de datos de fuentes de información como sistemas de información, correos electrónicos, redes sociales, sensores de internet de las cosas; en diferentes formatos o sin formato alguno, a esto lo llamamos información estructurada, semiestructura y no estructuradas. Las bases de datos no relacionales llamadas NOSQL de la sigla “**NOT ONLY SQL**”, es un sistema manejador de base de datos para trabajos de BIGDATA orientada a colecciones y documentos, utilizando para los documentos el formato JSON (Java Script Object Notation).

BSON es un formato de intercambio de datos usado principalmente para su almacenamiento y transferencia en la base de datos MongoDB. Es una

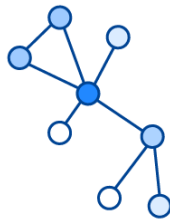
representación binaria de estructuras de datos y mapas. El nombre BSON está basado en el término JSON y significa Binary JSON (JSON Binario).<sup>i</sup>

## TIPO DE BASE DE DATOS NOSQL.

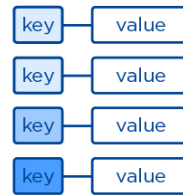
Document



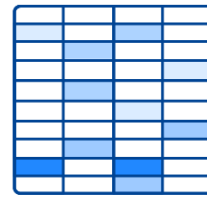
Graph



Key-Value



Wide-column



### ORIENTADAS A COLUMNAS

- Se almacena en columna como llave y valor
- Se encuentran en la misma partición
- Se incrementa la velocidad

### ORIENTADAS A DOCUMENTOS

- Almacena la información en forma de documentos
- Utiliza formatos semi estructurados como JSON
- Naturaleza de tipo jerárquica (documentos dentro de documentos)
- Una colección es un conjunto de documentos

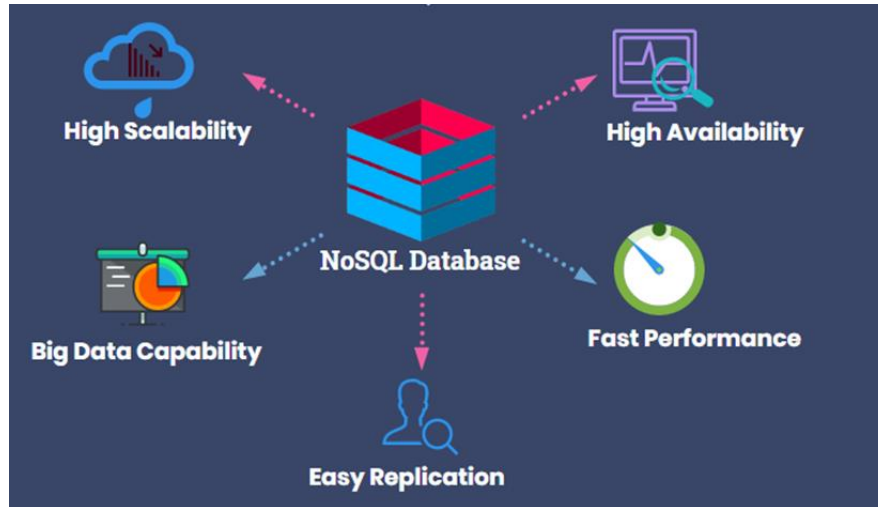
### ORIENTADOS POR CLAVE-VALOR

- Se almacena la información como clave-valor
- La clave es el identificador único
- Incrementa la velocidad de búsqueda
- Algunas se trabajan directamente en la memoria

### ORIENTADAS A GRAFOS

- Representan la información como nodos
- Representan las relaciones como aristas
- Estas almacenan información como nodos y aristas

## CARACTERISTICAS DE LAS BASES DE DATOS NOSQL



- Normalmente son distribuidas
- Escalables
- Flexibles
- No soportan subconsulta y uniones de tablas (JOIN)
- No soportan los métodos ACID (**A**tomicy, **C**onsistency, **I**solation, **D**urability)
- Garantiza el paradigma BASE (**B**ase **A**bility, **S**oft State, **E**ventual Consistency)
- Se almacena la información de forma flexible, donde un documento puede variar con respecto a las propiedades de otro documento de la colección
- Se utilizan para el manejo de grandes volúmenes de datos (BIGDATA)

## ESTRUCTURAS UTILIZADAS CON JSON

### COLUMNAS

- Identificador HASH dado por el SMDB-NOSQL (\_id)
- Propiedades Clave-Valor

```
{  
  "idemp": 1,  
  "nom": "Pedro Palitos",  
  "genero": "M",  
}
```



```
"edad": 23,  
"dni": 19407970  
}
```

## SUPER COLUMNNA

Documento compuesto por varias columnas.

```
{  
  "idemp":12,  
  "nom":"Carlos Villagran",  
  "edad":62,  
  "dni":{  
    "tipo":"CC",  
    "numero":19407906,  
    "fecha":"14/07/1979"  
  }  
}
```

## FAMILIA DE COLUMNAS

Un documento puede guardar otro documento.

```
{  
  "idemp":12,  
  "nom":"Carlos Villagran",  
  "edad":62,  
  "dni":{  
    "tipo":"CC",  
    "numero":19407906,  
    "fecha":"14/07/1979"  
  },  
  "cursos":["python","java","php","mongodb"],  
  "mensajes":[  
    {  
      "subject":"Reunión de planeación",  
      "body":"agenda de desarrollo web",  
      "from":"sotico@gmail.com",  
      "fecha":"2022-02-27"  
    }, {  
      "body":"Entrega de stickholder del proyecto",  
      "from":"cpinilla@gmail.com",  
      "fecha":"2022-03-12"  
    }  
  ]  
}
```





```
}  
]  
}
```

Antes de insertarlo se debe verificar que el JSON fue escrito de forma correcta, para esto se debe verificar con la pagina [Validador y formateador JSON Online - JSON Lint](#).

```
1 {  
2   "idemp": 1,  
3   "nom": "Pedro Palitos",  
4   "genero": "M",  
5   "edad": 23,  
6   "dni": 19407970  
7 }
```

Validar JSON Claro Soporte JSONLint por \$ 2 / mes

**Resultados**

Valid JSON



## PRACTICA MONGODB

Para comenzar la práctica, se debe descargar los archivos con GIT (ver “VIDEOS/Instalación/ clonarrepositoriogit.cmd” o sin GIT (ver “VIDEOS/Instalación/ descargar archivos mongodb.cmd” ). Los archivos se encuentran en: <http://github.com/fegasu/EDTMONGODB> o descargar [Aquí](#)

Descargar los siguientes archivos de instalación:

- Servidor MongoDB (SW/mongoserver.cmd)
- Consola MongoDB (SW/mongosh.cmd)
- Mongo Compas (SW/mongocompass.cmd)

## INSTALACION SERVIDOR MONGODB

Para la instalación del servidor MONGODB se recomienda ver el video: “VIDEOS/Instalación/mongoinstalacion.cmd”

## INSTALACION CONSOLA MONGODB

Para la instalación de la consola de MONGODB se recomienda ver el video: “VIDEOS/Instalación/instalarmongosh.cmd”

## MOSTRAR LAS BASE DE DATOS CREADAS

Ingresar a MONGOSH y digitar el comando:

```
show databases;
```

## CREAR UNA BASE DATOS

```
use prueba;
```

Tener en cuenta cuando se escriba db se refiere a la base de datos que se está usando. El comando use se debe utilizar para seleccionar una base de datos.

## CREANDO UNA COLECCIÓN

Una colección guarda un documento o muchos documentos, utilice el siguiente comando:

```
db.createCollection("emp")
```





## INSERTAR UN DOCUMENTO DENTRO DE UNA COLECCIÓN

Se utiliza la clausula “insertOne” de la siguiente manera para insertarlo dentro de la colección “emp”, para esto se debe escribir el documento en el formato JSON así:

```
emp1={
  "idemp": 1,
  "nom": "Pedro Palitos",
  "genero": "M",
  "edad": 23,
  "dni": 19407970
}
```

Para insertarlo dentro de la colección “emp” de la base de datos prueba

```
db.emp.insertOne(emp1);
```

Inserte los siguientes datos:

Idemp	Nom	Genero	Edad	dni
1	Pedro Palitos	M	23	19456970
2	Maria Casquito	F	32	
3	Celia Gacha	F	43	

## INSERTANDO MAS DE UN DOCUMENTO A LA VEZ

Construimos los documentos emp4 y emp5 así:

```
emp4={
  "idemp":4,
  "nom":"Maria la bandida",
  "genero":"F",
  "edad":32
}
emp5={
  "idemp": 5,
  "nom": "ROSA MELO",
  "genero": "F",
  "edad":25
}
```



Utilizamos la cláusula “insertMany” así:

```
db.emp.insertMany([emp4,emp5])
```

También podemos hacerlo así:

```
db.emp.insertMany([  
  {  
    "idemp":6,  
    "nom":"Juan Tolomeo",  
    "genero":"M",  
    "edad":37  
  },  
  {  
    "idemp": 7,  
    "nom": "Carlos Pichamata",  
    "genero": "M",  
    "edad":52  
  }  
])
```

Ingresar los siguientes empleados:

Idemp	Nom	Genero	Edad	dni
8	Manuela Beltran	F	15	123456
9	Monica Galindo	F	23	6789123
10	Elver Galindo	M	21	
11	Simón Tolomeo	M	51	1357901

## LISTAR TODOS LOS DOCUMENTOS DE UNA COLECCIÓN

Se utiliza la cláusula “find()” así:

```
db.emp.find();
```

Para buscar con una condición simple, por ejemplo, buscar el empleado 6:

```
db.emp.find(  
  {  
    idemp:"6"  
  }  
)
```

Para encontrar los empleados mayores a 30

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

Para encontrar los empleados mayores o iguales a 30

```
db.emp.find(  
  {edad:{$gte:30}})
```







Para encontrar los empleados menores a 30

```
db.emp.find(  
  {edad:{$lt:30}}  
)
```

Para encontrar los empleados menores o iguales a 30

```
db.emp.find(  
  {edad:{$lte:30}}  
)
```

Para encontrar los empleados de 23, 42 y 21

```
db.emp.find(  
  {  
    edad:{$in:[23,42,21]}  
  }  
)
```

## UTILIZANDO BUSQUEDAS CON OPERADORES LOGICOS

Utilizando el operador lógico AND

```
db.emp.find(  
  $and:[  
    {edad:{$lte:35}},  
    {genero:"F"}  
  ]  
)
```

Utilizando el operador lógico OR

```
db.emp.find(  
  $or:[  
    {edad:{$lte:35}},  
    {genero:"F"}  
  ]  
)
```

Encuentre los empleados que tienen "dni":

```
db.emp.find(  
  {  
    dni:{$exists:true}  
  }  
)
```

## ORDENANDO LA SALIDA DE LOS DOCUMENTOS

De forma descendente:

```
db.emp.find().sort(  
  {  
    edad:-1  
  })
```



Encontrar el empleado que tiene la máxima edad

```
db.emp.find().sort(  
  {  
    edad:-1  
  }  
).limit(1)
```

## UTILIZAR EXPRESIONES REGULARES

Encontrar empleados que comience con el patrón palabra "Simón"

```
db.emp.find(  
  {  
    nom:/^ Simón /  
  })
```

Encontrar empleados que contenga la palabra "ver"

```
db.emp.find(  
  {  
    nom:/ver/  
  }  
)
```

Cuantos documentos hay en la colección "emp"

```
db.emp.find().count()
```

7

Cuantos empleados hay con una edad menor o igual a 25

```
db.emp.find({edad:{$lte:25}}).count()
```

4

Mostrar dos documentos a partir del tercer documento

```
db.emp.find().skip(2).limit(2)
```

## MOSTRAR EL NOMBRE DE LOS DOCUMENTOS DE LA COLECCION

```
db.emp.find().forEach(df =>print(df.nom))
```

## ASIGNAR A UNA VARIABLE UN DOCUMENTO

```
var user=db.emp.findOne()  
user.nom='PEDRO PALOS GORDOS'
```

## ACTUALIZAR UN DOCUMENTO

```
db.emp.updateOne(  
  {idemp: 1} ,  
  {$set:{"nom":" Juan Charrasqueado"}}  
)
```





## ADICIONAR UN NUEVO ATRIBUTO A UN DOCUMENTO

```
db.emp.updateOne(  
  {idemp: 1} ,  
  {$set:{"Correo":"fegasu@misena.edu.co"}}  
)
```

## INCREMENTAR UN ATRIBUTO EN TODOS LOS DOCUMENTOS

```
db.emp.updateMany(  
  {} ,  
  {$inc:{edad:1}}  
)
```

## DECREMENTAR UN ATRIBUTO EN TODOS LOS DOCUMENTOS

```
db.emp.updateMany(  
  {} ,  
  {$inc:{edad:-1}}  
)
```

## BORRAR UN DOCUMENTO

```
db.emp.deleteOne(  
  {idemp:7}  
)
```

## ADICIONAR ATRIBUTOS A UN DOCUMENTOS

```
Emp12={  
  "idemp":12,  
  "nom":"Carlos Villagran",  
  "edad":62,  
  "dni":{  
    "tipo":"CC",  
    "numero":19407906,  
    "fecha":"14/07/1979"  
  },  
  "cursos":["python","java","php","mongodb"],  
  "mensajes":[  
    {  
      "subject":"Reunión de planeación",
```





```
"body":"agenda de desarrollo web",
"from":"sotico@gmail.com",
"fecha":"2022-02-27"
},
{
  "body":"Entrega de stickholder del proyecto",
  "from":"cpinilla@gmail.com",
  "fecha":"2022-03-12"
}
]
}
db.emp.insertOne(emp12)
```

## AGREGA LA PROPIEDAD A LOS DOCUMENTOS QUE NO LO TENGA.

Agrega la propiedad dni a los documentos que no lo tienen

```
db.emp.updateMany(
  {
    'dni':{$exists:false}
  },
  {
    $set:{'dni':{
      'tipo':'CC',
      'numero':0,
      'fecha':''
    }}
  }
)
```

## ADICIONA DOCUMENTOS AL FINAL.

```
db.emp.updateOne(
  {
    idemp:9
  },
  {$push:{
    mensajes:{
      subject:"Pruebas TYT",
      body:"Pruebas TyT Tecnologos Sena",
```



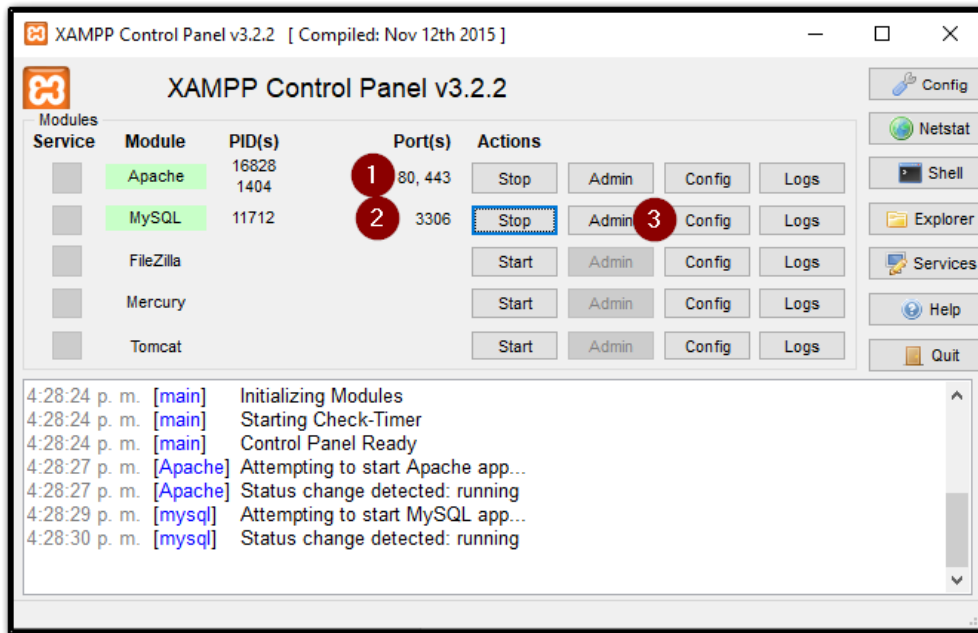


Servicio Nacional de Aprendizaje  
Formato Taller  
Centro de Gestión de Mercados, Logística y Tecnologías de la Información.

```
from:"secretaria@sena.edu.co",  
fecha:"2022-08-27"  
}  
}  
)
```

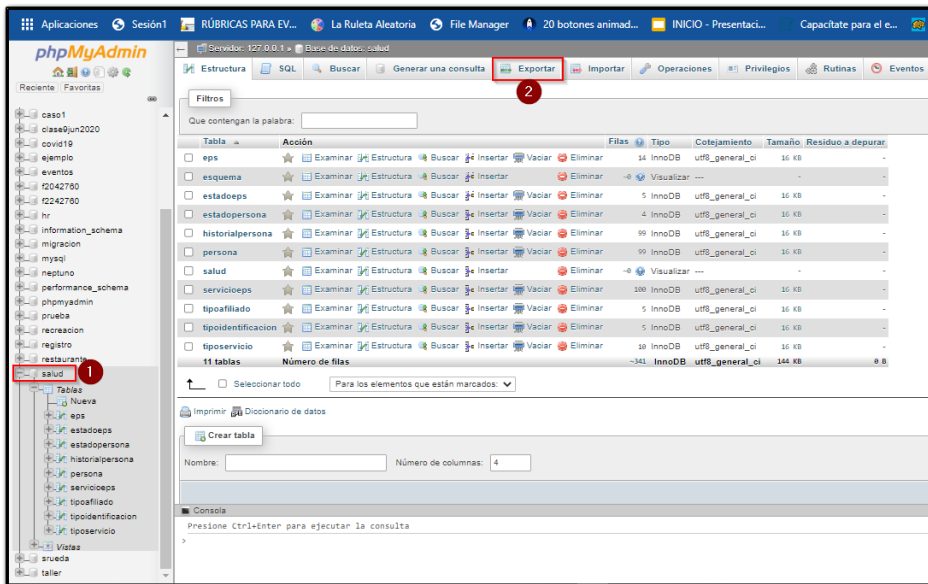
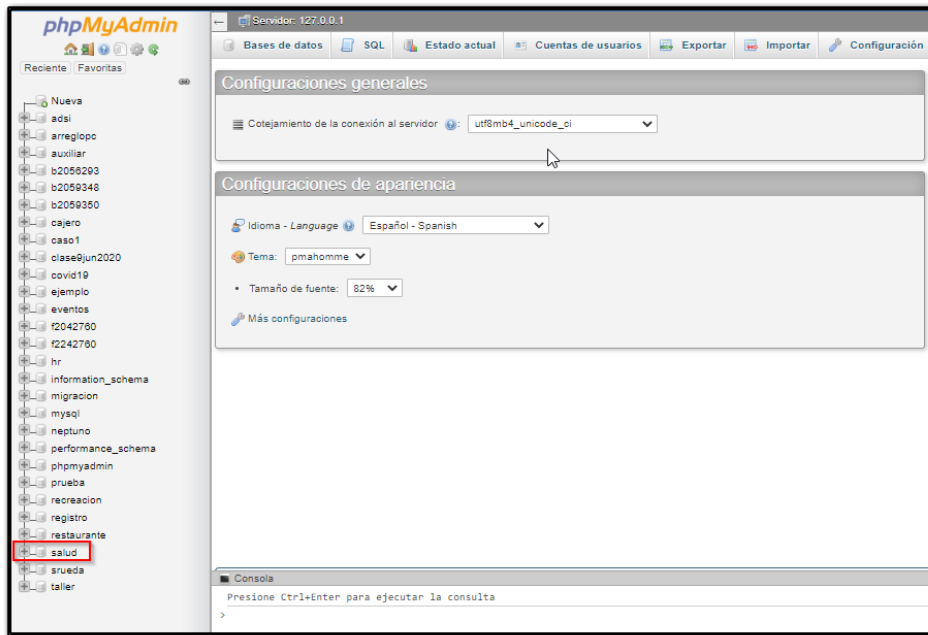


## MIGRAR BASE DE DATOS MYSQL A MONGODB



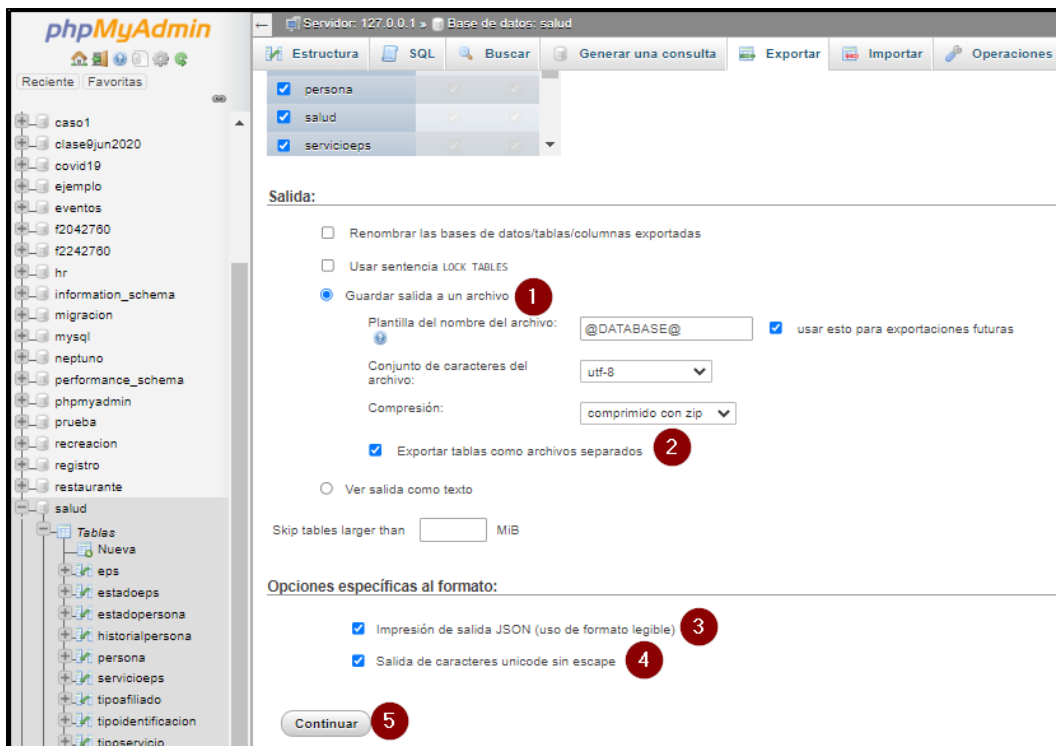


Servicio Nacional de Aprendizaje  
Formato Taller  
Centro de Gestión de Mercados, Logística y Tecnologías de la Información.





Servicio Nacional de Aprendizaje  
Formato Taller  
Centro de Gestión de Mercados, Logística y Tecnologías de la Información.







## CONVERTIR EXCEL A JSON

Utilizar la pagina <https://products.aspose.app/cells/es/conversion/excel-to-json>

**Ejercicio:** Convertir el archivo "DATOS/NEPTUNO/Neptuno.xls" y migrarlo a una base de datos MONGODB, se recomienda ver el video "[InsertManyNeptuno.mp4](#)"

## CONVIRTIENDO CSV A JSON

- Descargar el archivo "[COVID19-JULIO2020.zip](#)" que se encuentra en "DATOS/COVID19/COVID19-JULIO2020.zip"
- Descomprima
- Crear un archivo php llamado "CSV2JSON.php".

```
<meta http-equiv="Content-Type" content="text/html; charset=utf-8" />
<?php
$data = [];
$data=file("COVID19.csv");
$fp=fopen("COVID19.json","w");
$b=explode(';', $data[0]);
for($i=0;$i<count($b);$i++){
    if(strpos($b[$i], ' '))
        $b[$i]=str_replace(' ', '_', $b[$i]);
    elseif(is_integer($b[$i]))
        $b[$i]=preg_replace( "/\r|\n/", "", $b[$i] );
}
echo count($b);
for($i=1;$i<count($data);$i++){
    $c=explode(';', $data[$i]);
    $sal='db.casos.insertOne({';
    for($j=0;$j<count($c);$j++){
        if(is_int($c[$j]))
            $q=preg_replace( "/\r|\n/", "", $c[$j] );
        else
            $q=chr(39).preg_replace( "/\r|\n/", "", $c[$j] ).chr(39);
        $sal.=$b[$j].': '.$q;
        if($j<count($c)-1)
            $sal.=',';
    }
    $sal.='});';
    fwrite($fp,$sal.chr(10));
    echo $sal.chr(10);
}
```



Servicio Nacional de Aprendizaje  
Formato Taller  
Centro de Gestión de Mercados, Logística y Tecnologías de la Información.

```
fclose($fp);  
?>
```

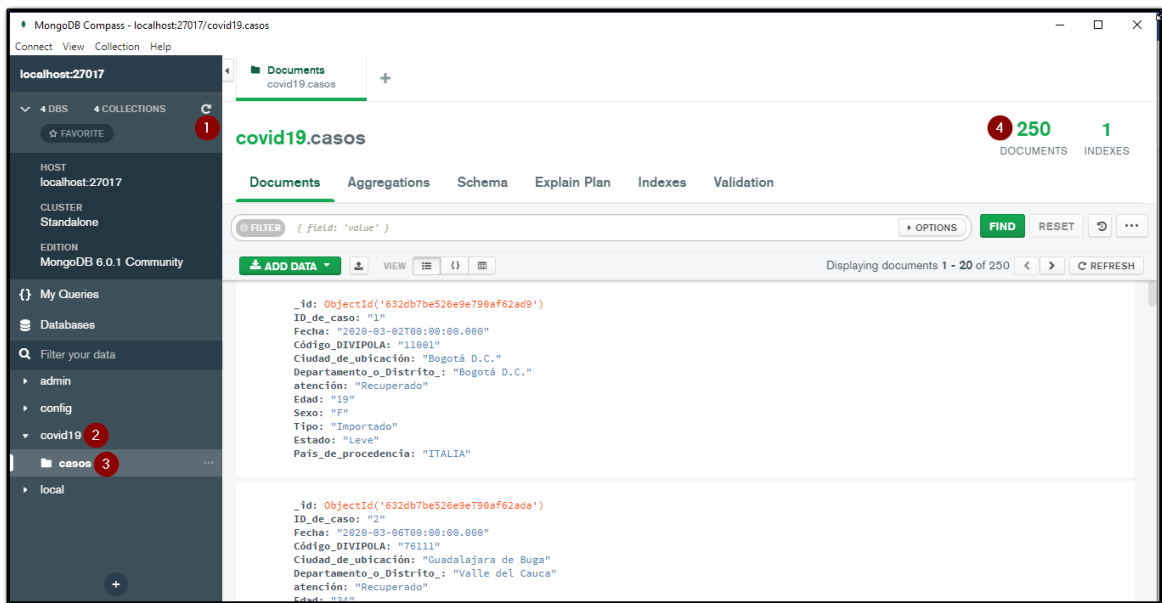
- Ejecutar el programa anterior donde lo creo, desde la línea de comando así:

```
D:\>php -f CSV2JSON.php
```

- Creará un archivo llamado "COVID19.json", ábralo, copie al portapapeles todo el contenido.
- Abra el programa MongoSh
- Elija la base de datos

```
test> use covid19
```

- Copiar del portapapeles, usando la tecla derecha del mouse.
- Ingrese al programa Mongo Compass y verifique el cargue





## CONECTANDO DESDE PYTHON A MONGODB

Se necesita instalar la biblioteca pymongo así:

Pip install pymongo

Realizar el siguiente código en un archivo llamado: "MongoCliente.py"

```
from pymongo import MongoClient
MONGO_URI='mongodb://localhost'
cliente=MongoClient(MONGO_URI)
db=cliente['prueba']
coleccion=db['emp']
eps1=coleccion.find()
for r in eps1:
    print(r['idemp']+"-"+r['nom'])
```

Ejecute el programa

Python MongoCliente.py

## CONECTANDO DESDE PHP A MONGODB<sup>ii</sup>

- Para descargar el controlador MongoDB para Windows, visita [pecl.php.net](http://pecl.php.net)
- Descargar el controlador [7.3 Thread Safe \(TS\) x86](#), este es para la versión de PHP 7.3.x y si se utiliza versión 8 o posterior utilice el drive correspondiente
- Descomprima el archivo descargado y copie los archivos "php\_mongodb.dll" y "php\_mongodb.pdb" en "C:\xampp\php\ext"
- Abrir el archivo "php.ini"
- Adicionar la línea:

extension=php\_mongodb.dll

- Reinicie de nuevo el apache
- Verifique con el phpinfo (<http://localhost/dashboard/phpinfo.php>) que xampp soporta MongoDB





#### mongodb

MongoDB support	enabled
MongoDB extension version	1.11.0
MongoDB extension stability	stable
libbson bundled version	1.10.1
libmongoc bundled version	1.10.1
libmongoc SSL	enabled
libmongoc SSL library	OpenSSL
libmongoc crypto	enabled
libmongoc crypto library	libcrypto
libmongoc crypto system profile	disabled
libmongoc SASL	enabled
libmongoc ICU	disabled
libmongoc compression	disabled
libmongocrypt bundled version	1.2.1
libmongocrypt crypto	enabled
libmongocrypt crypto library	libcrypto

Directive	Local Value	Master Value
mongodb.debug	no value	no value
mongodb.mock_service_id	Off	Off

- Escriba un archivo llamado “cnxMongoDb.php” en c:\xampp\htdocs”

```
<?php
echo "_____<br>";
$manager = new \MongoDB\Driver\Manager('mongodb://localhost:27017');
$options = ['limit' => 1];
$filter = ['_id' => new \MongoDB\BSON\ObjectId('63222b6f4c5aac0c8e499a9c')];
$query = new \MongoDB\Driver\Query($filter, $options);
echo "_____<br>";
$cursor = $manager->executeQuery('prueba.emp', $query);
$cursorArray = $cursor->toArray();
if(isset($cursorArray[0])) {
    var_dump($cursorArray[0]);
}
echo "_____<br>";
```

- Cambiar el ObjectId por uno \_id válido
- Regresa un arreglo BSON



## RETO POR DESARROLLAR

- Cargar los datos que se encuentran en “DATOS/COVID19/COVID19-JULIO2020.csv” en una base de datos MONGODB llamada “COVID” en una colección llamada “CASOS”.
- Realizar las siguientes consultas:
  - Cuantos casos se registraron en Bogotá
  - Cuantos casos se recuperaron en Bogotá
  - Cual fue la ciudad que más se recuperaron
  - Cual fue la ciudad que más se recuperaron siendo mujeres.
  - Cuantos fallecidos que fueron mujeres y mayores de 50 años.
  - Cuantos fallecidos que fueron hombres y menores de 50 años.
  - Cuantos casos fueron por personas del extranjero
  - Cuantos casos se registraron en Antioquía, menores de 30 años y por persona extranjeras
  - Cuantos casos se registraron por personas de España
  - Cuantos casos se registraron por personas de edad entre 20 y 50 años
- Realizar un programa que muestre cada una de las consultadas solicitadas en el punto anterior.

## PRODUCTOS SOLICITADOS.

- Documento de tipo texto con las consultas solicitadas
- Documento fuente del programa Python
- Informe grafico de desarrollo del taller.
- Ensayo sobre que son las bases de datos NOSQL y aplicaciones existentes.

## MODO DE ENTREGA.

Se debe empaquetar los productos solicitados en un archivo llamado “TallerNoSQL.zip”





Realizado por el instructor José Fernando Galindo Suarez

Centro de Gestión de Mercados, Logística y Tecnologías de la información

Jgalindos @sena.edu.co ©2022



<sup>i</sup> [que es bson - Búsqueda \(bing.com\)](https://www.bing.com)

<sup>ii</sup> [PHP => Utilizando MongoDB \(learntutorials.net\)](https://learntutorials.net)

redis



couchDB



neo4j



HBASE

riak



Cassandra

Sleek MongoDB