

# Progr. Python para Big Data

## Actividad 5: Funciones CRUD Pymongo y Studio3T.

### PASOS PREVIOS: CONEXIÓN BASE DE DATOS.

- Conectar nuestra base de datos con Studio3T.
- Crear la base de datos **actividad** y colección **notas**.

### CONEXIÓN CON BASE DE DATOS

- 1- Primero de todo, será necesario tener el **contenedor mongodb** de Docker activo:

<pre>\ docker ps</pre>							
CONTAINER ID	IMAGE	COMMAND	CREATED	STATUS	PORTS	NAMES	
76b12a935611	mongo	"docker-entrypoint.s..."	24 minutes ago	Up 24 minutes	0.0.0.0:27017->27017/tcp	mongodb	

**Nota:** Se puede ejecutar mediante `docker` o `docker compose`, en este caso se ha ejecutado con `docker-compose` con credenciales de acceso: `user admin`, `password admin`.

- 2- Una vez el contenedor activo, conectamos mongo a **Studio3T** teniendo en cuenta las credenciales y configuración necesaria:

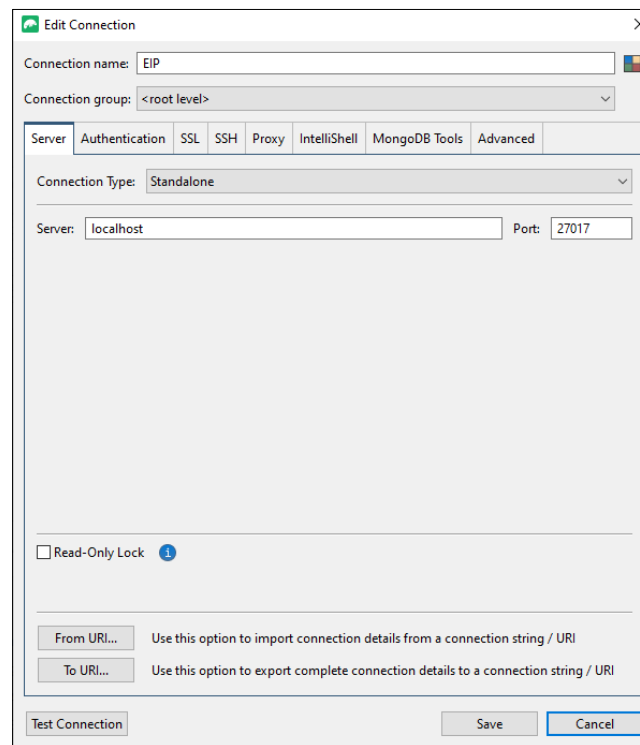
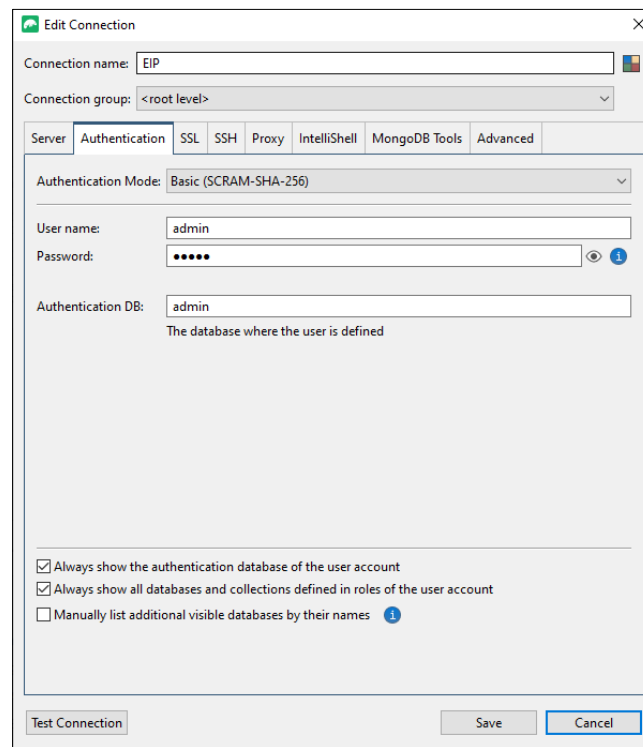
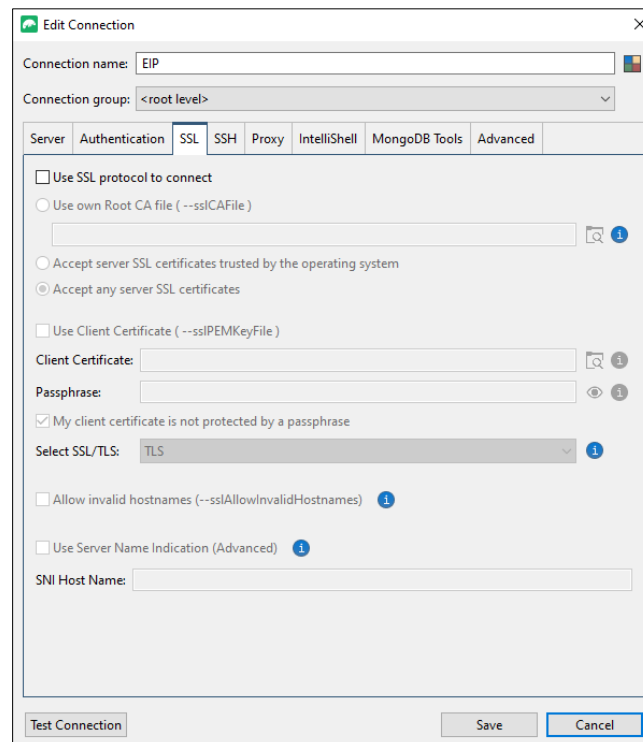


Ilustración 1: Nombre conexión y servidor mongodb



The screenshot shows the 'Edit Connection' dialog box with the 'Authentication' tab selected. The 'Connection name' is 'EIP' and the 'Connection group' is '<root level>'. The 'Authentication Mode' is 'Basic (SCRAM-SHA-256)'. The 'User name' is 'admin' and the 'Password' is masked with dots. The 'Authentication DB' is 'admin' with a note 'The database where the user is defined'. At the bottom, there are three checkboxes: 'Always show the authentication database of the user account' (checked), 'Always show all databases and collections defined in roles of the user account' (checked), and 'Manually list additional visible databases by their names' (unchecked). The 'Test Connection', 'Save', and 'Cancel' buttons are at the bottom.

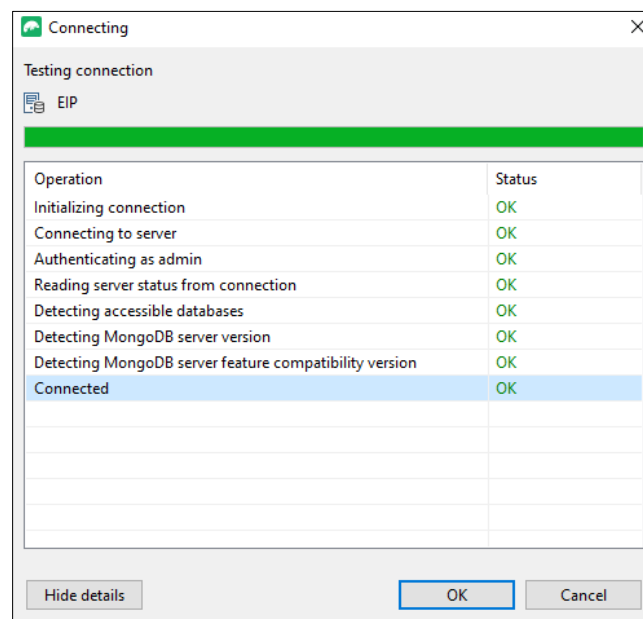
Ilustración 2: Credenciales acceso



The screenshot shows the 'Edit Connection' dialog box with the 'SSL' tab selected. The 'Connection name' is 'EIP' and the 'Connection group' is '<root level>'. The 'Use SSL protocol to connect' checkbox is unchecked. Below it, there are three radio buttons: 'Use own Root CA file ( --sslCAFile )' (unchecked), 'Accept server SSL certificates trusted by the operating system' (unchecked), and 'Accept any server SSL certificates' (selected). Below these, there are two checkboxes: 'Use Client Certificate ( --sslPEMKeyFile )' (unchecked) and 'My client certificate is not protected by a passphrase' (checked). The 'Client Certificate' and 'Passphrase' fields are empty. The 'Select SSL/TLS' dropdown is set to 'TLS'. At the bottom, there are two checkboxes: 'Allow invalid hostnames ( --sslAllowInvalidHostnames )' (unchecked) and 'Use Server Name Indication (Advanced)' (unchecked). The 'SNI Host Name' field is empty. The 'Test Connection', 'Save', and 'Cancel' buttons are at the bottom.

Ilustración 3: Protocolo SSL desactivado

3- Finalmente, testeamos que la conexión es correcta:



## CREACION BASE DE DATOS

- 1- Crear la base de datos **actividad** con la colección **notas**. Este paso, se puede realizar desde la interfaz de Studio3T, mediante código de programación o consola de mongo. En este caso se ha realizado mediante comandos de consola:
  - a. Acceder a terminal: `docker exec -it container_name bash`
  - b. Acceder a MongoDB Shell: `mongo -u user -p password`

```
root@76b12a935611:/# mongo -u admin -p admin
MongoDB shell version v5.0.9
connecting to: mongodb://127.0.0.1:27017/?compressors=disabled&gssapiServiceName=mongodb
Implicit session: session { "id" : UUID("38ef5006-36bb-4178-9deb-55b327a72ce5") }
MongoDB server version: 5.0.9
=====
Warning: the "mongo" shell has been superseded by "mongosh",
which delivers improved usability and compatibility. The "mongo" shell has been deprecated and will be removed in
an upcoming release.
For installation instructions, see
https://docs.mongodb.com/mongodb-shell/install/
=====
---
The server generated these startup warnings when booting:
  2022-07-06T16:07:49.485+00:00: Using the XFS filesystem is strongly recommended with the WiredTiger storage engine
  2022-07-06T16:07:50.581+00:00: /sys/kernel/mm/transparent_hugepage/enabled is 'always'. We suggest setting it to
  'never'
---
---
Enable MongoDB's free cloud-based monitoring service, which will then receive and display
metrics about your deployment (disk utilization, CPU, operation statistics, etc).

The monitoring data will be available on a MongoDB website with a unique URL accessible to you
and anyone you share the URL with. MongoDB may use this information to make product
improvements and to suggest MongoDB products and deployment options to you.

To enable free monitoring, run the following command: db.enableFreeMonitoring()
To permanently disable this reminder, run the following command: db.disableFreeMonitoring()
---
```

- 2- Una vez dentro de la Shell de mongo, crear la base de datos y colección para la actividad:
  - a. Crear base de datos y acceder a ella: use **db\_name**
  - b. Crear colección: `db.createCollection("collection_name")`

```
> show collections
> db.createCollection("notas")
{ "ok" : 1 }
> show collections
notas
> |
```

## EJERCICIO 1: INSERT

- Insertar el conjunto de documentos descritos en el manual de la práctica:

notas > _id					
_id	name	age	email	grade	date
62c5d08e17376987ca5618bb	Pedro López	25	pedro@eip.com	5.2	2022-07-06T20:12:30.806Z
62c5d08e17376987ca5618bc	Julia García	22	julia@eip.com	7.3	2022-07-06T20:12:30.806Z
62c5d08e17376987ca5618bd	Amparo Mayoral	28	amparo@eip.com	8.4	2022-07-06T20:12:30.806Z
62c5d08e17376987ca5618be	Juan Martínez	30	juan@eip.com	6.8	2022-07-06T20:12:30.806Z

Ilustración 4: Documentos de colección "notas"

Para realizar de forma correcta la adición de los valores del campo **date** en formato fecha, se ha recurrido a la validación de documentos. Esto permite mediante un **esquema json** validar que los datos que se introduce son los correctos, en este caso, un campo **date**.

The screenshot shows the 'Edit Validator' window in MongoDB. The breadcrumb path is 'admin > EIP (admin@l...lhost:27017) > actividad > notas'. The 'Validation Level' is set to 'strict' and the 'Validation Action' is set to 'error'. The instruction says 'Enter the validator document for this collection:'. A JSON schema is entered in the text area:

```
{
  "$jsonSchema": {
    "properties": {
      "name": {
        "bsonType": "string"
      },
      "age": {
        "bsonType": "int"
      },
      "email": {
        "bsonType": "string"
      },
      "grade": {
        "bsonType": "double"
      },
      "date": {
        "bsonType": "date"
      }
    }
  }
}
```

At the bottom, there are buttons for 'Validate JSON', 'Save', and 'Cancel'.

Ilustración 5: Validación de campos

```
INSERT error:
batch op errors occurred, full error: {'writeErrors': [{'index': 0, 'code':
rName': 'properties', 'propertiesNotSatisfied': [{'propertyName': 'date', 'd
]]}], 'errmsg': 'Document failed validation', 'op': {'name': 'Pedro López',
pserted': 0, 'nMatched': 0, 'nModified': 0, 'nRemoved': 0, 'upserted': []}]}
```

Ilustración 6: Insert datos no válidos

## EJERCICIO 2: UPDATE

- Actualizar los datos de los documentos indicados en el manual de la práctica:

























notas						
_id	name	age	email	grade	date	
 62c5d08e17376987ca5618bb	 Pedro López	 25	 pedro@eip.com	 5.2	 2022-07-06T20:12:30.806Z	
 62c5d08e17376987ca5618bc	 Julia García	 22	 julia@eip.com	 7.3	 2022-07-06T20:12:30.806Z	
 62c5d08e17376987ca5618bd	 Amparo Mayoral	 28	 amparo@eip.com	 9.3	 2022-07-06T20:12:30.806Z	
 62c5d08e17376987ca5618be	 Juan Martínez	 30	 juan@eip.com	 7.2	 2022-07-06T20:12:30.806Z	

Ilustración 7: UPDATE Amparo y Juan

## EJERCICIO 3: READ

- Mostrar por CMD los datos de la colección:

```
Total de documentos en actividad.notas: 4

{'_id': ObjectId('62c5d08e17376987ca5618bb'),
 'age': 25,
 'date': datetime.datetime(2022, 7, 6, 20, 12, 30, 806000),
 'email': 'pedro@eip.com',
 'grade': 5.2,
 'name': 'Pedro López'}
{'_id': ObjectId('62c5d08e17376987ca5618bc'),
 'age': 22,
 'date': datetime.datetime(2022, 7, 6, 20, 12, 30, 806000),
 'email': 'julia@eip.com',
 'grade': 7.3,
 'name': 'Julia García'}
{'_id': ObjectId('62c5d08e17376987ca5618bd'),
 'age': 28,
 'date': datetime.datetime(2022, 7, 6, 20, 12, 30, 806000),
 'email': 'amparo@eip.com',
 'grade': 9.3,
 'name': 'Amparo Mayoral'}
{'_id': ObjectId('62c5d08e17376987ca5618be'),
 'age': 30,
 'date': datetime.datetime(2022, 7, 6, 20, 12, 30, 806000),
 'email': 'juan@eip.com',
 'grade': 7.2,
 'name': 'Juan Martínez'}
```

Ilustración 8: Listado de documentos de la colección

EJERCICIO 4: BUSQUEDA

- Buscar los documentos cuyas notas se encuentren entre 7 y 7.5:

```
Documentos para la busqueda: 2

{'_id': ObjectId('62c5d08e17376987ca5618bc'),
 'age': 22,
 'date': datetime.datetime(2022, 7, 6, 20, 12, 30, 806000),
 'email': 'julia@eip.com',
 'grade': 7.3,
 'name': 'Julia García'}
{'_id': ObjectId('62c5d08e17376987ca5618be'),
 'age': 30,
 'date': datetime.datetime(2022, 7, 6, 20, 12, 30, 806000),
 'email': 'juan@eip.com',
 'grade': 7.2,
 'name': 'Juan Martínez'}
```

Ilustración 9: Resultado búsqueda en Python

The screenshot shows the Studio3T interface with a MongoDB connection to localhost:27017. A query is entered in the Shell: `db.getCollection("notas").find({'grade': {'$gte': 7.0, '$lte': 7.5}})`. The results are displayed in a table with columns: `_id`, `name`, `age`, `email`, `grade`, and `date`. Two documents are returned, corresponding to Julia García and Juan Martínez.

_id	name	age	email	grade	date
62c5d08e17376987ca5618bc	Julia García	22	julia@eip.com	7.3	2022-07-06T20:12:30.806Z
62c5d08e17376987ca5618be	Juan Martínez	30	juan@eip.com	7.2	2022-07-06T20:12:30.806Z

Ilustración 10: Resultado búsqueda Studio3T

EJERCICIO 5: DELETE

- Eliminar los datos del alumno Pedro López.



















notas > name					
_id	name	age	email	grade	date
 62c5d08e17376987ca5618bc	 Julia García	 22	 julia@eip.com	 7.3	 2022-07-06T20:12:30.806Z
 62c5d08e17376987ca5618bd	 Amparo Mayoral	 28	 amparo@eip.com	 9.3	 2022-07-06T20:12:30.806Z
 62c5d08e17376987ca5618be	 Juan Martínez	 30	 juan@eip.com	 7.2	 2022-07-06T20:12:30.806Z

Ilustración 11: Datos de Pedro eliminados