



CURSO: MINERÍA DE DATOS
MAESTRÍA EN EXPLOTACIÓN DE DATOS Y DESCUBRIMIENTO DE CONOCIMIENTO

LABORATORIO V: Introducción a Bases de Datos NOSQL

INTRODUCCIÓN

Esta práctica de laboratorio tiene como objetivo iniciar al estudiante en la gestión de datos mediante Bases de Datos NOSQL. Se hará hincapié en las operaciones CRUD –creación, lectura, actualización & eliminación de datos- y la consulta de datos. Para la resolución de estas consignas se utilizará MongoDB a efectos de poner en práctica los conceptos introducidos en clase.

CONSIGNAS

1. **OPERACIONES CRUD -Create, Read, Update, Delete-.** El objetivo de esta consigna es generar una DB con los artículos de tapa del Diario Página12¹.
 - a. Incorpore² en la colección *Noticias* de la DB *Pagina12*, los elementos textuales³ de los primeros 5 artículos que aparecen en tapa caracterizando a que sección pertenece cada una (política, educación, policiales, etc).
 - b. Ahora, verifique que los 5 documentos están guardados en la DB⁴.
 - c. Modifique⁵ el título de la última noticia ingresada en la DB por el siguiente “NOTICIA MODIFICADA EN DMUBA”.
 - d. Por último, elimine⁶ la 3 noticia de las que están guardadas en la DB.

2. **CONSULTAS**
 - a. Junto con estas consignas, dispone de un dataset con tweets en español que contienen la palabra “dólar”. Incorpore estos tweets en una DB Mongo⁷.
 - b. Seleccione⁸ todos los documentos que poseen al menos un retweet.

¹ <https://www.pagina12.com.ar/>

² Inspeccione los métodos *insertOne()* e *insertMany()* en la documentación de MongoDB.

³ Puede verificar los elementos de una noticia en el siguiente enlace:

https://es.wikipedia.org/wiki/Noticia#Elementos_de_una_noticia

⁴ Consulte el método *find()* en la documentación de referencia.

⁵ Puede hacerlo con los métodos *update()*, *updateOne()*, *updateMany()* y *replaceOne()*.

⁶ Inspeccione los métodos *deleteMany()* y *deleteOne()*.

⁷ Puede incorporarlo con el siguiente comando:

`mongoimport --db nombre_nueva_base_datos --collection nombre_coleccion --file path_donde_esta_json --jsonArray`



CURSO: MINERÍA DE DATOS

MAESTRÍA EN EXPLOTACIÓN DE DATOS Y DESCUBRIMIENTO DE CONOCIMIENTO

- c. Ahora, seleccione todos los documentos cuyo usuario comience con la letra “P”.

3. MONGODB & R

- a. R dispone de varias librerías⁹ para gestionar datos mediante Bases de Datos MongoDB. Elija alguna de ellas e indague como realizar las operaciones de los puntos 1 y 2 desde R a través de esas librerías.

Referencias sugeridas:

- MongoDB Documentation: <https://docs.mongodb.com/manual/>
- McCreary, D., & Kelly, A. (2013). Making Sense of NoSQL: A guide for managers and the rest of us.
- Documentación ‘mongolite’ R Link-1:
<https://cran.r-project.org/web/packages/mongolite/mongolite.pdf>
- Documentación ‘mongolite’ R Link-2: <https://jeroen.github.io/mongolite/>

⁸ Verifique el método *find()* para utilizarlo aplicando criterios de búsqueda.

⁹ Entre ellas, las librerías más conocidas son *mongolite* (**recomendada**), *RMongo* y *rmongodb*.