



**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 avanzada de datos.

Para la resolución de estas consignas se utilizará la Base de Datos Mongo DB, la cual está orientada a documentos, 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 Página12<sup>1</sup>.
  - a. Incorpore<sup>2</sup> en la colección *Noticias* de la DB *Pagina12*, los elementos textuales<sup>3</sup> 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 DB<sup>4</sup>.
  - c. Modifique<sup>5</sup> el título de la última noticia ingresada en la DB por el siguiente “NOTICIA MODIFICADA EN DMUBA”.
  - d. Por último, elimine<sup>6</sup> 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 orientada a documentos.
  - b. Seleccione<sup>7</sup> todos los documentos que poseen al menos un retweet.

---

<sup>1</sup> <https://www.pagina12.com.ar/>

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

<sup>3</sup> Puede verificar los elementos de una noticia en el siguiente enlace:

[https://es.wikipedia.org/wiki/Noticia#Elementos\\_de\\_una\\_noticia](https://es.wikipedia.org/wiki/Noticia#Elementos_de_una_noticia)

<sup>4</sup> Consulte el método *find()* en la documentación de referencia.

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

<sup>6</sup> Inspeccione los métodos *deleteMany()* y *deleteOne()*.



**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”.
- d. MongoDB posee la capacidad de definir índices para realizar de manera más eficiente la búsqueda de texto. Indague<sup>8</sup> sobre la gestión de estos índices y la búsqueda de texto.

**3. MONGODB & R**

- a. R dispone de varias librerías<sup>9</sup> 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/>

---

<sup>7</sup> Verifique el método *find()* para utilizarlo aplicando criterios de búsqueda.

<sup>8</sup> Indague la documentación de referencia en <https://docs.mongodb.com/manual/text-search/>

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