

TC3041, Bases de Datos Avanzadas
Mini Proyecto MongoDB
Actividad en EQUIPO

Fechas de Entrega:

Objetivo:

Conocer de primera mano un sistema de bases de datos NoSQL. En este caso MongoDB. Aprender a usar las herramientas que proporciona MongoDB para importar datos y conocer las distintas maneras de explotar información por medio de queries en MongoDB. Integrar MongoDB con Tableau para visualizar la información obtenida de manera gráfica.

Instrucciones

Utilizaremos la base de datos para [COVID19](#) que proporciona la Secretaría de Salud.

En la liga encontrará 2 documentos la Bases de Datos y un Diccionario de Datos. El diccionario de datos contiene información acerca de qué significa cada ID, por ejemplo el catálogo de TIPO_PACIENTE , 1 - Ambulatorio, 2- Hospitalizado, etc. La base de datos es el registro como tal del paciente.

- 1) La base de datos está bien hasta cierto punto normalizada. Es decir, tenemos diferentes “tablas” conectadas por medio de IDs, como por ejemplo TIPO_PACIENTE. El primer paso que tenemos que hacer es **de-normalizarla**; es decir, reemplazar esos IDs con el valor actual. Si tenemos. TIPO_PACIENTE 1 en la tabla principal, tenemos que reemplazar el “1” con el valor del catálogo, “AMBULATORIO”.
- 2) Una vez que tenemos la base de datos no normalizada, en un solo CSV. Podemos usar una [herramienta](#) de MongoDB para importarla hacia nuestro modelo
- 3) Cada integrante del equipo tiene que construir 3 queries de complejidad avanzada que ayuden a aportar información importante acerca de la pandemia en México.
- 4) Por último conectaremos nuestro modelo de mongo hacia Tableau para visualizar los queries del punto número 3. [Aquí](#) encontrará información acerca del conector. Considerar al menos un query de tipo “time-series” para ver el comportamiento de la pandemia a través del tiempo y visualizarlo con Tableau.

Para el punto #1 puede utilizar las herramienta de su preferencia. Una opción podría ser crear un script para leer el csv y crear uno nuevo con las características principales. Por ejemplo puede usar [python](#) para un script sencillo que reemplace los IDs con valores actuales.

Para el punto #3, considere que el query se va a visualizar en Tableau en el punto 4. Hacer el query lo más informativo posible para que la visualización sea justificada. Por ejemplo, datos generales del país por día o por mes, datos estadísticos por estado, un Pie Chart con los datos de confirmados, sospechosos, etc. Hay mucha información, ¡ **Sean creativos !**

Entregables

- . Código de cada uno de los queries.
- . Screenshot de cada uno de los queries en Mongo y también visualizados en Tableau.
- . Reflexiones personales de cada uno de los integrantes del equipo.
- . **No** es necesario construir una interfaz gráfica.