Carrera: INGENIERIA EN INFORMATICA

Asignatura: 3641 – Bases de Datos Aplicada.

Tema: NoSQL

Unidad: Unidad 4

Objetivo: Comprender los conceptos básicos del modelo BASE. Que el alumno sea capaz de implementar en la práctica las instrucciones básicas del lenguaje NoSql a través de la herramienta MongoDB

Competencia/s a desarrollar:

- Identificación, formulación y resolución de problemas de ingeniería en sistemas de información/informática.
- Concepción, diseño y desarrollo de proyectos de ingeniería en sistemas de información / informática.
- Gestión, planificación, ejecución y control de proyectos de ingeniería en sistemas de información / informática.
- Utilización de técnicas y herramientas de aplicación en la ingeniería en sistemas de información / informática.
- Generación de desarrollos tecnológicos y/o innovaciones tecnológicas.
- Desempeño en equipos de trabajo.
- Comunicación efectiva.

Descripción de la actividad:

- 1. Tiempo estimado de resolución: 1 semana
- 2. Metodología: En computadora.
- 3. Forma de entrega: No obligatoria
- 4. Metodología de corrección y feedback al alumno: Presencial y por Miel.

Contenido

Guía práctica	3
Creación de Base de datos	3
Creación de Colecciones	3
Inserción de Documentos	5
Tipos de datos	9
Actualización de datos	12
Borrado de documentos	17
Importar Documentos (CVS)	19
Indices	20
Vistas	23
Operaciones de agregación	28
Practica adicional	34
Documentación complementaria	34

Guía práctica

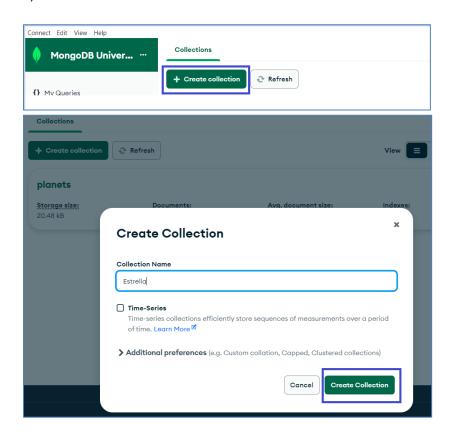
Ejercicios

Creación de Base de datos

- 1) Instalar la Herramienta Mongo DB (ver ppt). En caso de tener inconvenientes o dudas comunicarse por los canales indicados.
 - a) Conectarse a la base de datos On Cloud indicada en la ppt
 - b) Crear una Base de datos local

Creación de Colecciones

- 2) Colección: Creación de la colección
 - a) A través de la GUI



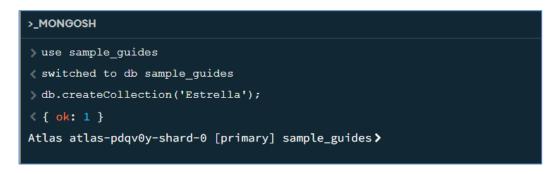
Presionar Refresh

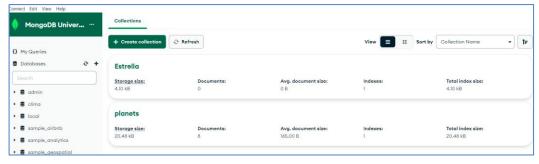


b) Por línea de comando:

Por defecto, mongosh, se conecta a la base de datos test. Para usar una base de datos diferente, ejecute el siguiente comando en mongosh: use <database name>

use sample_guides
db.createCollection('Estrella');





Para ver las colecciones por linea de comando:

db.getCollectionNames()

Inserción de Documentos

- 3) Insertar documentos
 - a) Por la GUI



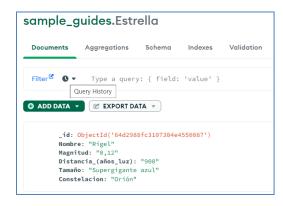
Elegir la opción Insert Document

```
* Paste one or more documents here
*/

{
  "Nombre": "Rigel",
  "Magnitud": "0,12",
  "Distancia_(años_luz)": "900",
  "Tamaño": "Supergigante azul",
  "Constelacion": "Orión"
```



Presionar el botón Insert



b) Por línea de comando

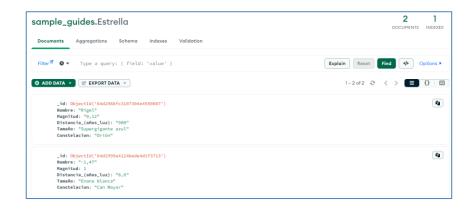
```
db.Estrella.insertOne({
  "Nombre": "-1,47",
  "Magnitud": 1,
  "Distancia_(años_luz)": "8,6",
  "Tamaño":"Enana blanca",
  "Constelacion": "Can Mayor"
})
```

```
>_MONGOSH

> db.Estrella.insertOne({
    "Nombre": "-1,47",
    "Magnitud": 1,
    "Distancia_(años_luz)": "8,6",
    "Tamaño":"Enana blanca",
    "Constelacion": "Can Mayor"
})

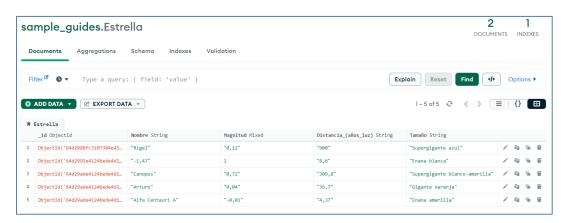
< {
    acknowledged: true,
    insertedId: ObjectId("64d2995e4124bede4d1f3713")
}

Atlas atlas-pdqv0y-shard-0 [primary] sample_guides>
```



```
4) Insertar/Agregar varios documentos por línea de comando:
    db.Estrella.insertMany([
     {
             "Nombre": "Canopus",
              "Magnitud": "0,72",
              "Distancia_(años_luz)": "309,8",
             "Tamaño": "Supergigante blanco-amarilla",
              "Constelacion": "Carina"
     },
     {
             "Nombre": "Arturo",
             "Magnitud": "0,04",
             "Distancia_(años_luz)": "36,7",
              "Tamaño": "Gigante naranja",
              "Constelacion": "Boyero"
     },
              "Nombre": "Alfa Centauri A",
             "Magnitud": "-0,01",
             "Distancia_(años_luz)": "4,37",
              "Tamaño": "Enana amarilla",
              "Constelacion": "Centauro"
     }
    ])
    Atlas atlas-pdqv0y-shard-0 [primary] sample_guides>> db.Estrella.insertMany([
                                                 "Tamaño": "Supergigante blanco-amarilla",
                                                 "Nombre": "Arturo",
                                                 "Magnitud": "0,04",
                                                 "Tamaño": "Gigante naranja",
                                                 "Constelacion": "Boyero"
                                                 "Nombre": "Alfa Centauri A",
                                                 "Magnitud": "-0,01",
                                                 "Tamaño":"Enana amarilla",
```

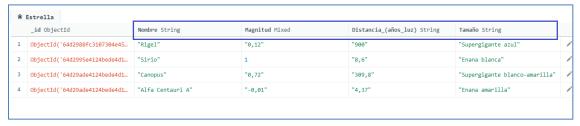
```
> MONGOSH
      "Constelacion": "Carina"
      "Nombre": "Arturo",
     "Magnitud": "0,04",
     "Distancia_(años_luz)": "36,7",
      "Tamaño": "Gigante naranja",
      "Constelacion": "Boyero"
      "Nombre": "Alfa Centauri A",
     "Magnitud": "-0,01",
     "Distancia (años luz)": "4,37",
      "Tamaño":"Enana amarilla",
      "Constelacion": "Centauro"
< {
    acknowledged: true,
     '0': ObjectId("64d29ade4124bede4d1f3714"),
      '1': ObjectId("64d29ade4124bede4d1f3715"),
      '2': ObjectId("64d29ade4124bede4d1f3716")
Atlas atlas-pdqv0y-shard-0 [primary] sample_guides>
```



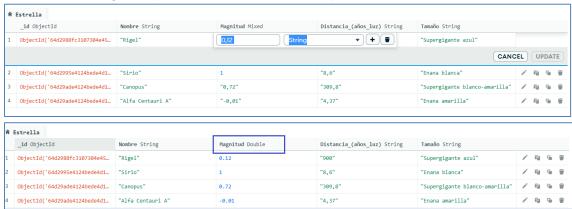
Tipos de datos

5) Tipos de datos

Al crear un documento con datos el tipo se dato se asigna según como se setea el contenido al momento del insert



Desde la GUI podemos modificar el tipo de dato que queremos establecer para ese campo. Al corregir el tipo de dato, se redefine el tipo de dato de la colección.



Recuerden que las Base de datos NoSQL nos ofrecen flexibilidad, es decir, en cada campo se pueden de cada documento se pueden colocar diferentes tipos de datos.

Crear la colección Asteroide

Insertar los documentos según la tabla que se encuentra debajo Modificar el tipo de dato en caso de que sea necesario.

Nombre	Dimensiones	Año Descubrimiento	
Winchester 747	171.71 km	1913	
Hilda 153	170.63 km	1875	
Pretoria 790	170 km	1912	
Aegle 96	169.9 km	1868	

String

Este tipo de datos es uno de los más utilizados en MongoDB para almacenar datos, los strings BSON son de tipo UTF-8. Por ende, los controladores para cada lenguaje de

programación se convierten del formato de string del lenguaje a UTF-8 mientras serializan

y desrealizan BSON. El string debe ser un UTF-8 válido.

Integer

Los datos de tipo integer (entero) en MongoDB se usan para almacenar valores numéricos. Podemos almacenar el tipo de datos integer (entero) de 2 formas: entero de

32 bits y entero de 64 bits con signo.

Double

Este tipo de datos se utiliza para almacenar los valores de punto flotante, tener en cuenta

que el separador decimal es el punto '.' veamos el ejemplo a continuación:

"Distancia_(años_luz)": 8.6

"Magnitud": 0.71

Booleano

Los booleans o booleanos se utilizan para almacenar valores true o false, veamos un

ejemplo a continuación:

"hasRings": false

"hasRings": true

Null

Este tipo de dato se utiliza para definir el valor de un campo como nulo o inexistente.

"Constelacion": null

Array

Este tipo de datos esta compuesto por un conjunto de valores en su interior, en un array podemos almacenar valores del mismo tipo en una sola clave, veamos un ejemplo a

continuación:

"mainAtmosphere": ["CO2", "N"]

Object (Objeto)

Este tipo de datos almacena documentos incrustados, a los cuales también se les conoce como documentos anidados. Los documentos incrustados o anidados son aquellos tipos de documentos que contienen un documento dentro de otro documento. Veamos un ejemplo a continuación:

```
"surfaceTemperatureC": {
    "min": null,
    "max": null,
    "mean": 464
}
```

Date (Fecha)

Este tipo de datos es también uno de los más usados, pues se considera que la mayoría de datos debe tener una fecha para poder gestionarlos mejor. El tipo de dato fecha es un entero (integer) de 64 bits que representa el número de milisegundos. El tipo de datos BSON generalmente admite la fecha y hora UTC y está firmado. Si el valor de este tipo de datos es negativos, entonces representa fechas anteriores al año 1970.

Existen diferentes métodos para devolver fechas, se puede devolver como un string o como un objeto de fecha usando Date(), new Date() y new ISODate().

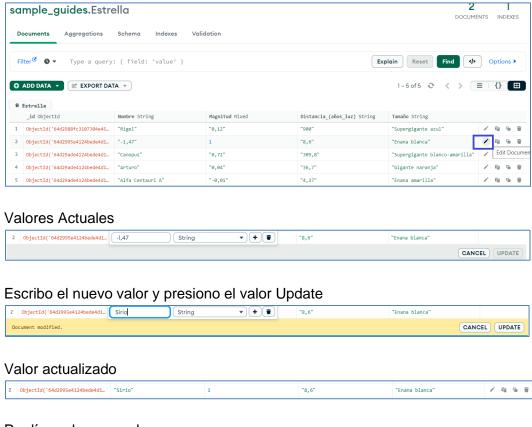
```
"Fecha Descubrimiento": "Wed Oct 16 2012 09:16:11 GMT-0500 "
```

"Fecha Descubrimiento": "2012-10-16T00:00:02.012+00:00"

"Fecha Descubrimiento": "2012-10-16"

Actualización de datos

- 6) Update/Actualizar
 - a) Por la GUI



b) Por línea de comando:

db.Estrella.update({Nombre:"Rigel"},{\$set:{"Fecha Descubrimiento":null}})

```
>.MONGOSH

Atlas atlas-pdqv0y-shard-0 [primary] sample_guides > db.Estrella.update()(Nombre: "Rigel"), ($set: {"Fecha Descubrimiento":null}))
```

Se verifica que se realizó el cambio observando los valores que se encuentran en los parámetros:

```
acknowledged: true, insertedId: null, --> valor insertado matchedCount: 1, --> cantidad de documentos encontrados modifiedCount: 0,--> cantidad de documentos modificados upsertedCount: 0 --> cantidad de documentos creados
```

```
>_MONGOSH

> db.Estrella.update({Nombre:"Rigel"},{$set:("Fecha Descubrimiento":null}})

< DeprecationWarning: Collection.update() is deprecated. Use updateOne, updateMany, or bulkWrite.

<{
    acknowledged: true,
    insertedId: null,
    matchedCount: 1,
    modifiedCount: 1,
    upsertedCount: 0
}

Atlas atlas-pdqv0y-shard-0 [primary] sample_guides>
```

Otro ejemplo

db.Estrella.update({_id:"64d2995e4124bede4d1f3713"},{\$set:{"Fecha Descubrimiento":null}})

```
>,MONGOSH

Atlas atlas-pdqv0y-shard-0 [primary] sample_guides≯db.Estrella.updatd{{_id:"64d2995e4124bede4d1f3713"},{$set:{"Fecha Descubrimiento":null}}}
```

En este caso vemos que no encontró el objeto indicado, por lo tanto, no realiza ninguna modificación.

```
>_MONGOSH

> db.Estrella.update({_id:"64d2995e4124bede4d1f3713"},{$set:{"Fecha Descubrimiento":null}})

< {
    acknowledged: true,
    insertedId: null,
    matchedCount: 0,
    modifiedCount: 0,
    upsertedCount: 0
}

Atlas atlas-pdqv0y-shard-0 [primary] sample_guides>
```

El error se debe a la forma en que se indica el campo query, para poder realizar la búsqueda por ese campo se debe escribir de la siguiente manera:

```
db.Estrella.update(
{_id: ObjectId("64d2995e4124bede4d1f3713")},
{
    $set: {
        "Fecha Descubrimiento":null
     }
})
```

Y vemos en la gui el valor actualizado



Update y UpdateOne funcionan de la misma manera, he aquí un ejemplo: db.Estrella.updateOne(

```
{Nombre: "Canopus"},
{
    $set: {
        "Fecha Descubrimiento":null
    }
})
```

7) Armar el script para realizar un insert, utilizando los datos que se encuentran debajo.

Nombre	Magnitud	Distancia (años luz)	Tamaño	Constelación
Sirio	-1,47	8,6	Enana blanca	Can Mayor
Canopus	0,72	309,8	Supergigante blanco- amarilla	Carina
Arturo	0,04	36,7	Gigante naranja	Boyero
Alfa Centauri A	-0,01	4,37	Enana amarilla	Centauro
Vega	0,03	25,3	Gigante blanca	Lira
Rigel	0,12	900	Supergigante azul	Orión
Procyon	0,34	11,4	Supergigante azul	Canis Minor
Achernar	0,5	144	Gigante blanca	Eridano
Betelgeuse	0,3-1,2	640	Supergigante roja	Orión
Hadar	0,6	530	Gigante blanca	Centauro
Capella	0,71	42	Gigante amarilla	Auriga
Altair	0,8	16,8	Enana blanca	Águila
Aldebarán	0,9	65	Gigante naranja	Tauro
Spica	1,4	260	Gigante azul	Virgo
Antares	1,09	550	Supergigante roja	Scorpio
Pollux	1,15	33,7	Gigante roja	Géminis
Fomalhaut	1,16	25	Enana-subgigante blanca	Piscis Austrinus
Deneb	1,34	1.425	Supergigante blanca	Cygnus
Becrux	1,3	280	Subgigante azul	Crux Australis
Alfa Centauri B	1,2	4,5	Enana naranja	Centaurus
Regulus	1,35	77	Subgigante azul	Leo
Acrux A	1,33	325	Subgigante azul	Crux Australis
Shaula	1,62	570	Subgigante azul	Scorpio
Gacrux	1,63	88	Gigante roja	Crux Australis
Bellatrix	1,64	240	Gigante azul	Orión
Elnath	1,68	131	Gigante azul	Tauro
Beta Carinae	1,7	110	Subgigante blanca	Carina
Alnilam	1,7	1.340	Supergigante azul	Orión
Alnitak A	1,7	700	Supergigante azul	Orión

8) Al dato insertado en el punto anterior realizar el update del campo Fecha Descubrimiento por el valor null

Borrado de documentos

- 9) Comando delete
 - a) Por la GUI



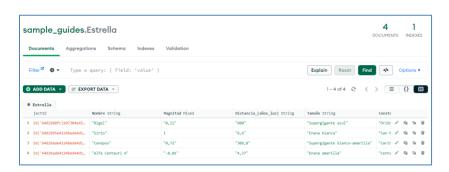
Se selecciona el documento que se desea eliminar y se presiona el botón delete



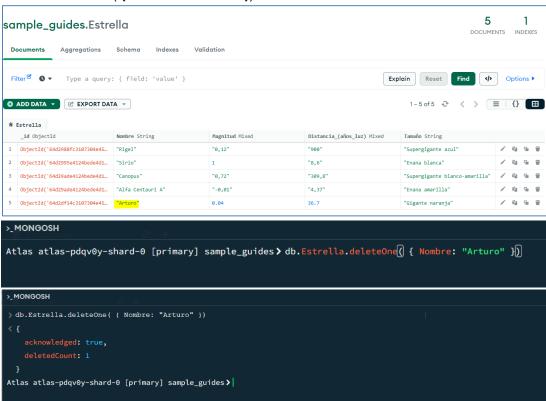
Se presiona el botón delete



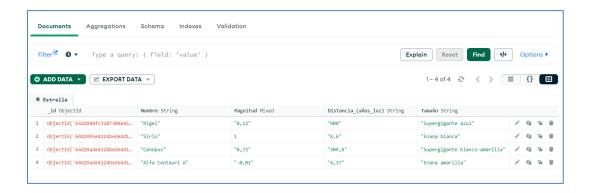
Se verifica el documento eliminado



b) Por línea de comando
 Seleccionamos el documento a eliminar
 db.movies.delete({ Nombre: "Arturo" })



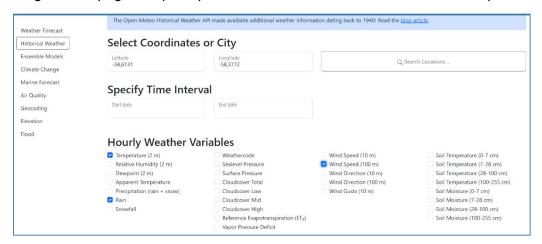
Se verifica que el documento no se encuentra en la base



Borrar uno de los documentos insertados

Importar Documentos (CVS)

10) Crear la base de datos Clima de forma local importando un archivo cvs Dirigirse a la pagina https://open-meteo.com/en/docs/historical-weather-api

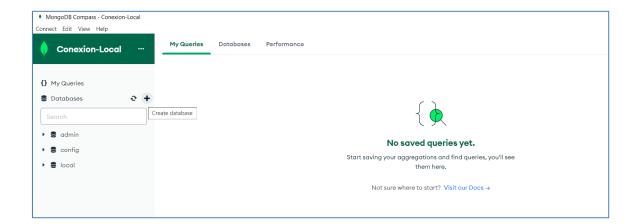


Elegir como locacion Buenos Aires

Seleccionar un rango de fechas de 6 meses

Descargar el archivo en formato CVS.

- a) Crear la base de datos llamada Clima (desde línea de comando o desde la GUI)
- b) Crear la colección llamada: clima_diario (desde línea de comando o desde la GUI)
- c) Desde la colección importar el archivo descargado.
- d) Observar los tipos de datos.
- e) Que sucede si se importa el archivo tal cual se descargó?
- f) Como son los tipos de datos importados?



Indices

11) Crear un índice a la colección clima diario.

Un índice posibilita el acceso directo y rápido haciendo más eficiente las búsquedas. Sin índice, MongoDB debe recorrer secuencialmente todos los documentos de una colección para encontrar un documento en particular.

En MongoDB el primer índice que se crea en forma automática corresponde al campo _id, recordemos que dicho campo podemos asignarle un valor nosotros o hacer que se cree automáticamente

Los índices son estructuras asociadas a las colecciones, una colección almacena los campos indexados y se crean para acelerar las consultas.

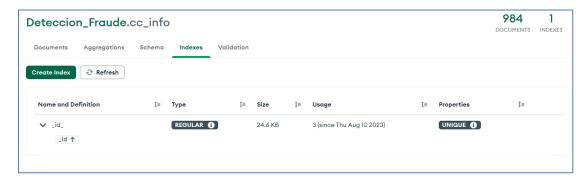
No se recomienda crear índices sobre campos que no se usan con frecuencia en consultas o en colecciones muy pequeñas.

Los cambios sobre las colecciones, como inserción, actualización o eliminación de documentos son incorporados automáticamente en los archivos índices.

La desventaja es que consume espacio en el disco y genera costo de mantenimiento (tiempo y recursos) cuando se efectúan inserciones y modificaciones.

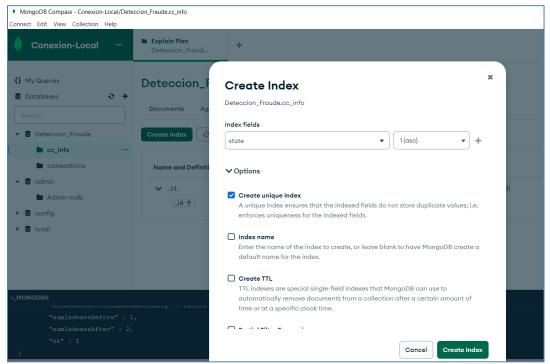
- Crear la base de datos Deteccion_Fraude de forma local
- Descomprimir el archivo archive.zip
- Crear una colección por cada archivo csv
- Importar el archivo

Vemos que por defecto se crea de forma automática el índice sobre el campo id



Para crear un índice

a) Por la GUI



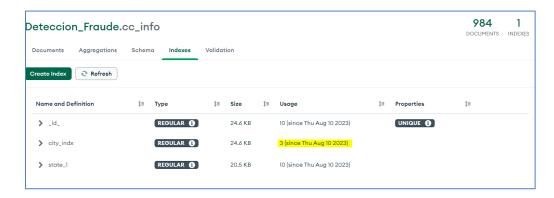
b) Por línea de comando
 Visualización de los índices existentes

db.cc_info.getIndexes()

db.cc_info.createIndex({"city_indx" : 1})
db.cc_info.getIndexes()

Uso del índice en búsquedas

db.cc_info.find({ city: "Dallas" })



Para borrar los índices: dropIndex() El método se usa para eliminar el índice único y el método dropIndexes () se usa para eliminar múltiples índices. db.cc_info.dropIndex("city_indx")

Vistas

```
12) Creación de vistas
   a) Por linea de comando
   db.createView( "Transacciones_Estado", "transactions", [
       $lookup:
           from: "cc_info",
           localField: "credit_card",
           foreignField: "credit card",
           as: "Transacciones Estado"
         }
     },
{
       $project:
                          _id: 0,
                         credit card:1,
                         date:1,
                         transaction dollar amount:1,
                         city: "$Transacciones_Estado.city"
                  }
       { $unwind: "$city"}
   ])
   $lookup
   Se utiliza para "unir" documentos cuandolos campos localField: y foreignField
   coinciden.
   Los documentos coincidentes se agregan como una matriz (as:)
   Es el equivalente al Join/where de sgl
   $project
   Se selecciona un subconjunto de los campos disponibles.
   Es el Equivalente al Select
```

_id: 0 → con cero indicamos que no vamos a utilizar ese campo (false) credit_card:1 → indifcamos que vamos a utilizar ese campo (true) city:"\$Transacciones_Estado.city"→ indicamos que es el campo de la tabla "foranea" por ende lo invocamos utilizando las comillas dobles y el signo pesos.

\$unwind

Convierte el campo de una matriz a un valor escalar.

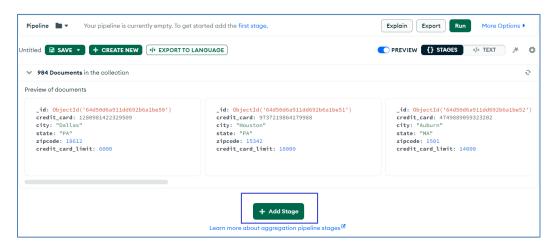
Al traer los campos de otro documento, si no se utiliza este parametro va a crear la vista con

teniendo los campos de la tabla en una matriz.

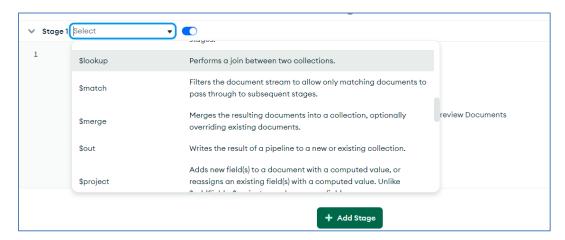
b) Por la GUI
 Se debe ingresar a la colección y desde alli ir a "Aggregations"



Click en Add Storage



Dentro del Stage 1 seleccionar \$lookup



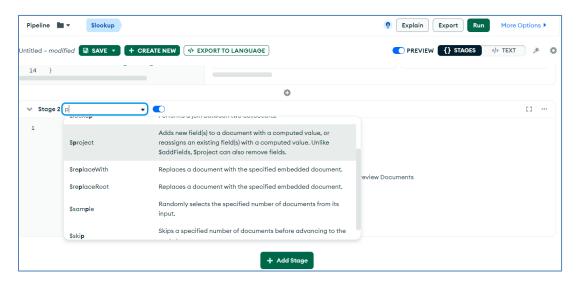
Completar los campos según corresponda con la vista a crear

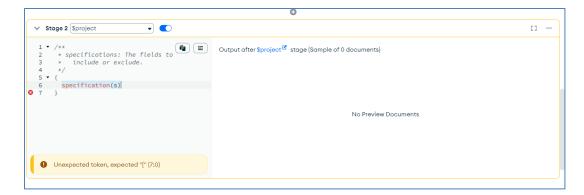


En este caso completo con los mismo campos de la vista creada anteriormente



Procedemos a agregar un nuevo Stage, en este caso elegimos project





Y completamos con los campos que deseamos incluir en la vista Nota: Cuando el campo no pertenece a la colección desde donde la estamos creando, el mismo se crea en forma de matriz; por ello se debe utilizar luego el parametro \$unwind

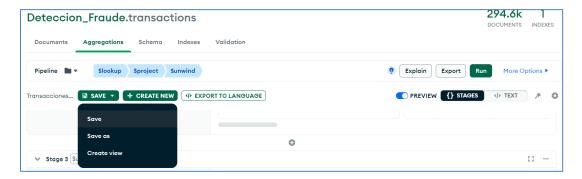


Agregamos un nuevo Stage y colocamos el campo city, tal como se muestra en la imagen

Ya se visualiza el campo city como un String y no como un array.



Por último guardamos la vista seleccionando Create View

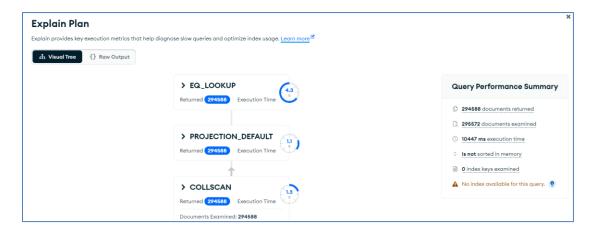




Visualizamos desde la base de datos la vista creada



Para ver el plan de ejecucion y las metricas del mismo, se presiona sobre el boton Explain y allí nos muestra las estadisticas.



Operaciones de agregación

13) Operaciones de agregación

Aggregation Pipeline - Canalización

Una canalización/tuberia de agregación consta de una o más etapas que procesan documentos, es la forma en que en mongodb se le da tratamiento a los documentos/colecciones.

Cada etapa (stage) realiza una operación sobre los documentos de entrada. Por ejemplo, una etapa puede filtrar documentos, agrupar documentos y calcular valores.

Los documentos que salen de una etapa pasan a la etapa siguiente.

Una canalización de agregación puede devolver resultados para grupos de documentos. Por ejemplo, devuelva los valores total, promedio, máximo y mínimo.

Las operaciones de agregación se utilizan mayoritariamente para:

- Agrupar valores de varios documentos.
- Procesamiento y operaciones matemáticas.
- Analizar cambios de datos a lo largo del tiempo.

Las operaciones aggregation pueden guardarse como consultas (querys) o guardarse como vistas.

Nota: utilizar la base Deteccion_Fraude para las siguientes actividades

Agrupar valores:

\$match: funciona de forma similar al método find() o el filter. Nos permite filtrar documentos de la colección según la consulta que pasemos como parámetro.

db.cc_info.aggregate([{\$match : {city: "Dallas" }}])

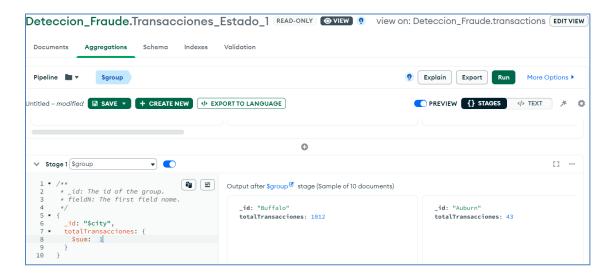
Procesamiento y operaciones matemáticas:

\$sum: Permite acumular los valores pasamos como parámetro Sumamos el total de gastos realizados por una tarjeta de credito, agrupados por estado origen de la tarjeta

b.Transacciones_Estado_1.aggregate([{\$group: {_id: "\$city", total: { \$sum: "\$transaction_dollar_amount" } }}])

O tambien se puede utilizar como contador: Contar la cantidad de transacciones por estado

Puede realizarse por la GUI, se muestra la siguiente imagen a modo de ejemplo:



\$max, \$min

Permite obtener el valor máximo y el valor mínimo de la expresión que se evalúa para cada documento generado por el agrupamiento

Se pueden combinar los parametros vistos, según la necesidad de obtener datos: Ejemplo: obtener la suma total de transacciones dado un estado db.Transacciones_Estado_1.aggregate([

```
{
    $match: { city: "Houston" }
},
{
    $group: { _id: "$city", total: { $sum: "$transaction_dollar_amount" } }
}
}

> db.Transacciones_Estado_1.aggregate([
    {
        $match: { city: "Houston" }
    },
    {
        $group: { _id: "$city", total: { $sum: "$transaction_dollar_amount" } }
}

}

class of the content of the city of
```

Operadores comparativos:

\$gt, \$gte, \$lt, \$lte:

Estos operadores nos permiten comparar desigualdades, \$gt y \$gte para "mayor que" y "mayor o igual que" respectivamente, y \$lt y \$lte para "menor que" y "menor o igual que" respectivamente.

Retornar true cuando se cumple el criterio y false en caso contrario

```
db.Transacciones_Estado_1.find({
   transaction_dollar_amount: { $gte: 500, $lte: 700 }
})
```

\$eq, \$ne:

Estos operadores nos permiten comparar la igualdad o diferencia de dos expresiones que le pasamos como parámetro.

\$eq retorna true cuando los valores son equivalentes y false cuando no lo son, \$ne hace lo opuesto.

db.Transacciones_Estado_1.find({ transaction_dollar_amount: { \$eq : 0.01 }})

\$cmp:

Toma dos valores en una matriz y devuelve un número entero. El valor devuelto es:

Un número negativo si el primer valor es menor que el segundo. Un número positivo si el primer valor es mayor que el segundo. O si los dos valores son iguales.

```
>_MONGOSH
> db.Transacciones_Estado_1.aggregate([
     {$match: { credit_card: 9383760568579492,
            date: "2015-08-16 19:02:57"
     {$project:
           {result: {$cmp:["transaction_dollar_amount", 43.78]}}
< {
     result: 1
Deteccion_Fraude>
             Smatch Sproject
                                                                                   itled - modified SAVE T + CREATE NEW ( EXPORT TO LANGUAGE
                                                                                     PIPELINE OUTPUT
                                                                                                          OUTPUT OPTIONS *
                                                               Sample of 1 document
       $match:
    /**
                                                                                                                     r<sub>a</sub>
           * query: The query in MQL.
                                                                 result: 1
           credit_card: 9383760568579492,
date: "2015-08-16 19:02:57",
        $project:
13
14 •
15
16
17
18 •
19 •
20 •
21
22
23
24
          * specifications: The fields to
           result: {
    $cmp: [
    "transaction_dollar_amount",
```

Practica adicional

Usando la base clima

- Ingresar el campo Mes, en el cual se indique el nombre del mes según la fecha indicada en el documento
- Ingresar el campo ciudad, colocarle el dato Buenos Aires para todos los documentos
- Descargar un nuevo archivo con los mismos campos, y el mismo rango de fechas para la localidad Montevideo
- Crear una vista con el promedio de los campos temperature_2m (°C), rain (mm), windspeed 100m (km/h) por dia
- Obtener el máximo y el mínimo de temperatura diaria promedio
- Obtener la/s fechas en los que se produjo temperatura diaria promedio
- Obtener para un dia determinado (fecha a eleccion) las condiciones meteologicas de Buenos Aires y Montevideo

Documentación complementaria

14) Documentación de los comandos disponibles en Mongo DB

https://www.mongodb.com/docs/manual/reference/limits/#std-label-restrictions-on-db-names

https://www.mongodb.com/docs/mongodb-shell/crud/#std-label-mdb-shell-crud

https://www.mongodb.com/docs/manual/reference/method/

https://www.mongodb.com/docs/manual/aggregation/

https://www.youtube.com/watch?v=6weTUQVtiD8