

Cloud Computing: Servicios y Aplicaciones



UNIVERSIDAD DE GRANADA

PRÁCTICA 3: MONGODB

Autor

Juan Manuel Castillo Nievas



MÁSTER PROFESIONAL EN INGENIERÍA INFORMÁTICA 2020-2021

Granada, 21 de mayo de 2021

Índice

1. Preparación del entorno	2
2. Consultas	4
2.1. Totalizar el número de delitos por cada tipo	4
2.2. Franja horaria (o franjas) con más número de delitos	5

1. Preparación del entorno

Para la preparación de los datos en **MongoDB** se ha creado una base de datos en **MongoDB Atlas** (aprovechando que ya se tiene el usuario creado, el cluster creado y demás para el TFM) y he utilizado **MongoDB Compass** para importar los datos desde el csv. También he aprovechado el shell que ofrece MongoDB Compass para hacer las consultas correspondientes.

En la Figura 1 se muestra la creación de la base de datos llamada **practica3** con la creación de la colección llamada **datos**.

En la Figura 2 se muestra la importación del fichero *Sacramentocrime.January2006.csv* a la base de datos creada usando la interfaz gráfica de MongoDB Compass.

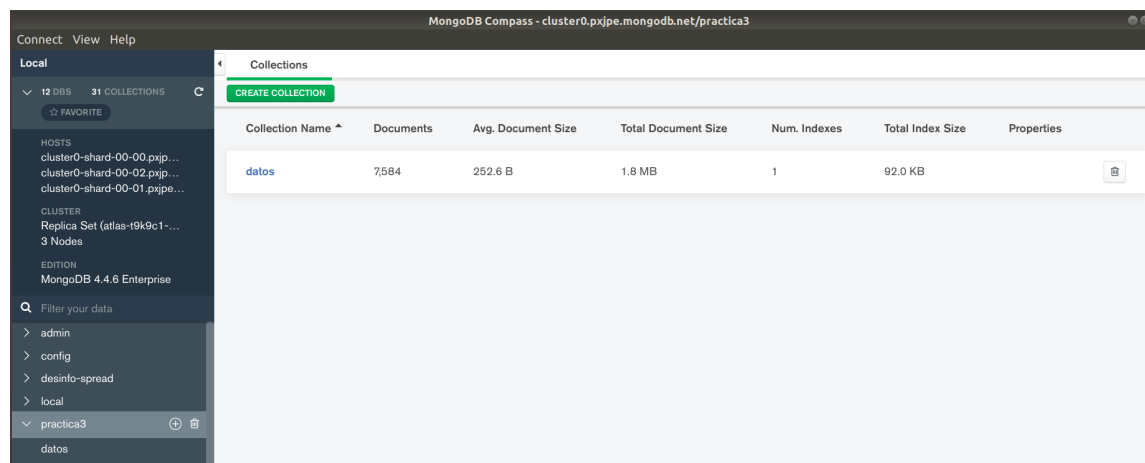


Figura 1: Creación de la base de datos y colección en MongoDB Compass

MongoDB Compass - cluster0.pxjpe.mongodb.net/practica3.datos

Import To Collection practica3.datos

Select File

/home/jumacasi/Documentos/master_ingenieria_informatica/CCSA/practi [BROWSE](#)

Select Input File Type

JSON CSV

Options

Select delimiter COMMA

☒ Ignore empty strings

☐ Stop on errors

Specify Fields and Types

	<input checked="" type="checkbox"/> cdatetime	<input checked="" type="checkbox"/> address	<input checked="" type="checkbox"/> district	<input checked="" type="checkbox"/> beat	<input checked="" type="checkbox"/> grid
	String	String	String	String	Strir
1	1/1/06 0:00	3108 OCCIDENTAL DR	3	3C	1115
2	1/1/06 0:00	2082 EXPEDITION WAY	5	5A	1512
3	1/1/06 0:00	4 PALEN CT	2	2A	212
4	1/1/06 0:00	22 BECKFORD CT	6	6C	1443
5	1/1/06 0:00	3421 AUBURN BLVD	2	2A	508
6	1/1/06 0:00	5301 BONNIEMAE WAY	6	6B	1084
7	1/1/06 0:00	2217 16TH AVE	4	4A	957
8	1/1/06 0:00	3547 P ST	3	3C	853
9	1/1/06 0:00	3421 AUBURN BLVD	2	2A	508
10	1/1/06 0:00	1326 HELMSMAN WAY	1	1B	444

CANCEL IMPORT

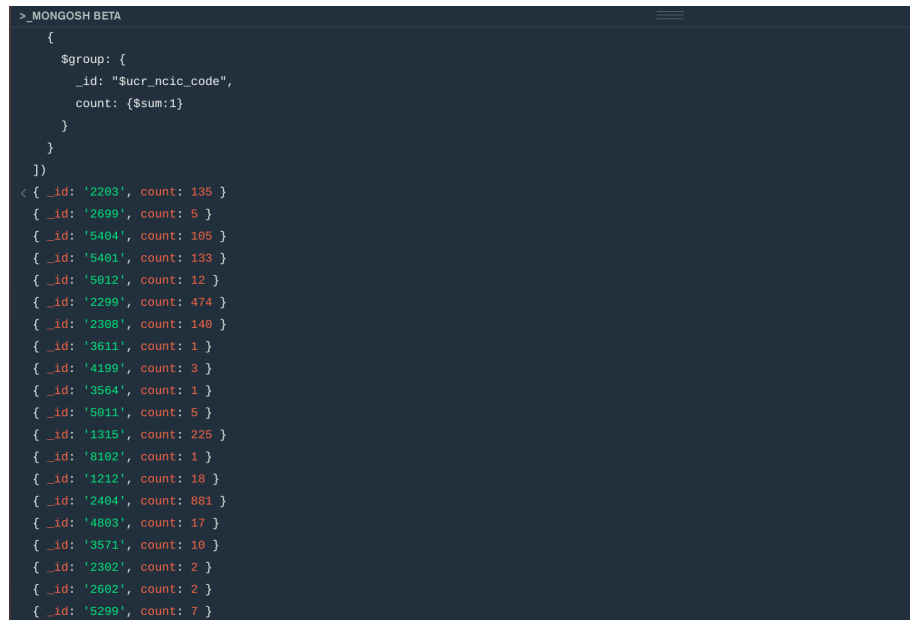
Figura 2: Importación del csv en MongoDB Compass

2. Consultas

2.1. Totalizar el número de delitos por cada tipo

```
1 db.datos.aggregate([  
2   {  
3     $group: {  
4       _id: "$ucr_ncic_code",  
5       count: {$sum:1}  
6     }  
7   }  
8 ])
```

En la Figura 3 se muestra la ejecución de la consulta en MongoDB Compass y la correspondiente salida.



The screenshot shows a terminal window titled "> _MONGOSH BETA". It displays the execution of the aggregation query from the previous block. The output is a list of 20 documents, each representing a crime type and its count. The documents are as follows:

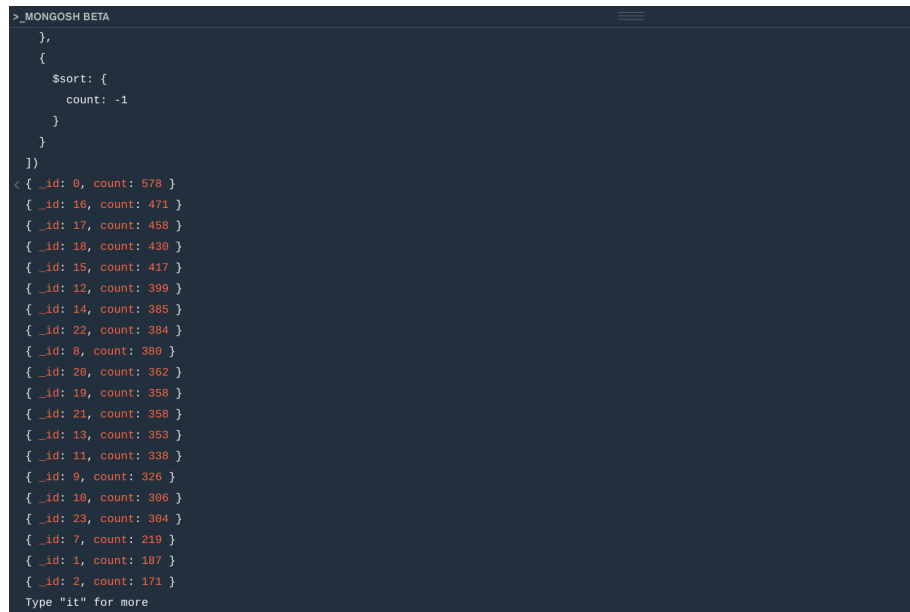
_id	count
'2203'	135
'2699'	5
'5404'	105
'5401'	133
'5012'	12
'2299'	474
'2308'	140
'3611'	1
'4199'	3
'3564'	1
'5011'	5
'1315'	225
'8102'	1
'1212'	18
'2404'	881
'4803'	17
'3571'	10
'2302'	2
'2602'	2
'5299'	7

Figura 3: Primera consulta: totalizar el número de delitos por cada tipo

2.2. Franja horaria (o franjas) con más número de delitos

```
1 db.datos.aggregate([
2   {
3     $group: {
4       _id: {
5         $hour:{
6           $dateFromString: {dateString: "$cdatetime", format: "%m/%d/%Y %H:%M"}
7         }
8       },
9       count: {$sum:1}
10    }
11  },
12  {
13    $sort: {
14      count: -1
15    }
16  }
17 ])
```

En la Figura 4 se muestra la ejecución de la consulta en MongoDB Compass y la correspondiente salida. La franja horaria con más delitos es la franja que empieza a las **00:00** con **578 delitos**.



```
>_MONGOSH BETA
},
{
  $sort: {
    count: -1
  }
}
])
{ _id: 0, count: 578 }
{ _id: 16, count: 471 }
{ _id: 17, count: 458 }
{ _id: 18, count: 430 }
{ _id: 15, count: 417 }
{ _id: 12, count: 399 }
{ _id: 14, count: 385 }
{ _id: 22, count: 384 }
{ _id: 8, count: 380 }
{ _id: 20, count: 362 }
{ _id: 19, count: 358 }
{ _id: 21, count: 358 }
{ _id: 13, count: 353 }
{ _id: 11, count: 338 }
{ _id: 9, count: 326 }
{ _id: 10, count: 306 }
{ _id: 23, count: 304 }
{ _id: 7, count: 219 }
{ _id: 1, count: 187 }
{ _id: 2, count: 171 }
Type "it" for more
```

Figura 4: Segunda consulta: franja horaria (o franjas) con más número de delitos