

Universidad Internacional de La Rioja Facultad de Ingeniería y Tecnología

Máster Universitario en Análisis y Visualización de Datos Masivos / Visual Analitics & Big Data

Uso Avanzado de BD NoSQL

Actividad de estudio presentado por:	Juan David Escobar Escobar.
	Andrés Felipe Leal Mora.
	Juan Manuel Bautista Correa.
	William Ramiro Ríos Henao.
Tipo de trabajo:	Actividad 3
Modalidad:	Individual / Grupal
Profesor/a:	Marlon Cárdenas
Fecha:	Marzo 2022

# Índice de contenidos

	ster Universitario en Análisis y Visualización de Datos Masivos / Visual Analitics & Data	
1.	Inicia la monitorización en MongoDB	3
2.	Inicia la monitorización en MongoDB	4
3.	Exportar la base de datos	4
4.	Restaurar una base de datos	5
5.	Caso de uso	5
6.	Conversion CSV a JSON	6
7.	Consultas y actualizaciones en MongoDB	10
8.	Neo4j	16
9.	Anexos	19

# 1. Inicia la monitorización en MongoDB

- 1. Inicia la monitorización en MongoDB, ejecuta las siguientes instrucciones:
  - a. En una consola diferente, visualiza las estadísticas con la instrucción
  - mongostat --discover

```
> Top 5 Operaciones
(base) juandavidescobarescobar@MacBook-Pro-de-JUAN Actividad 3 % mongostat --discover
    host insert query update delete getmore command dirty used flushes vsize res qrw arw net_in
                                                                                                    Estadisticas Mongo DB
net_out conn
                            time
             ÷α
                   ÷α
                          +a
                                *0
localhost
                                               010 0.0% 0.0%
                                                                    0 5.31G 38.0M 0|0 1|0
  47.8k
           3 Feb 16 09:29:28.502
localhost
          *0 *0 *0
3 Feb 16 09:29:29.496
                                               1|0 0.0% 0.0%
                                                                    0 5.31G 38.0M 0|0 1|0
  48.5k
localhost
                                               0|0 0.0% 0.0%
                                                                    0 5.31G 38.0M 0|0 1|0
           *0 *0 *0 *0
```

b. En otra consola distinta, visualiza el *top* de las operaciones contra la base de datos, utiliza la instrucción mongotop 5 donde 5 es el tiempo en segundos que debe refrescar la visualización.



c. Tened estas dos consolas siempre activas y visibles mientras realizáis los siguientes puntos. Interpretad su salida y describid brevemente qué muestran cuando se ejecutan dichas instrucciones (utiliza capturas para ilustrar tu respuesta).

**Interpretación:** La operación mongostat –discover, reporta un log de estadísticas de todos los miembros replica set o sharded del cluster de MongoDB. La ventana de monitoreo registra en el log del total de registros insertados, seleccionados, actualizados o eliminados, también permite evaluar el uso de la cantidad de memoria utilizada en GB, el tamaño de entrada y salida de red utilizado en MB, y el porcentaje de algunos recursos de maquina utilizados.

# 2. Inicia la monitorización en MongoDB

### 2. Importar la base de datos:

a. Cread la base de datos *inspections* y dentro de ella la colección que corresponda a los datos del fichero *act-grupal-city\_inspections.json*.

b. Cread la base de datos *countries* y dentro de ella dos colecciones que correspondan con los ficheros *act-grupal-countries-small.json* y act-grupal-countries-big.json.

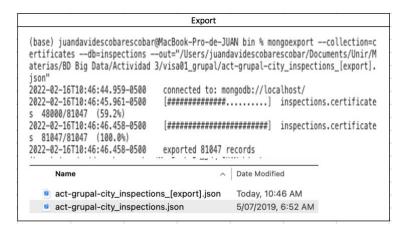
```
SMALL-COUNTIERS

> db.small.findOne()
{
    "_id": ObjectId("55a0f42f20a4d760b5fc3064"),
    "altSpellings": [
    "A0"
},
    "area": 14000000,
    "borders": [],
    "callingCode": [],
    "capital": "",
    "cca2": "A0",
    "cca3": "ATA",
    "cca3": "ATA",
    "cio": "",
    "demonym": "Antarctican",
    "landLocked": false,
    "landlocke
```

# **3.** Exportar la base de datos

### 3. Exportar la base de datos:

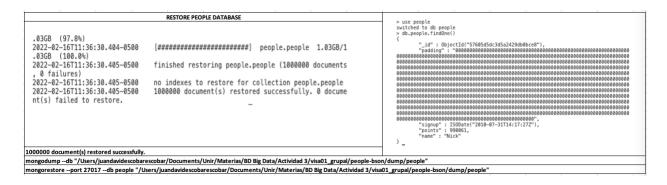
a. Haced un export de la colección inspections. Llamad al fichero de salida inspections.json.



### 4. Restaurar una base de datos

#### 4. Restaurar una base de datos:

- a. Para restaurar una base de datos, debéis indicar la ruta del directorio *dump* donde se almacenan los ficheros de la base de datos *(json y bson)*.
  - dump
    - people
    - --people.bson
    - --people.metadata.json
    - b. Descomprimid el fichero act-grupal-people.zip y ubicad los ficheros en la estructura de directorios indicada en el paso anterior.



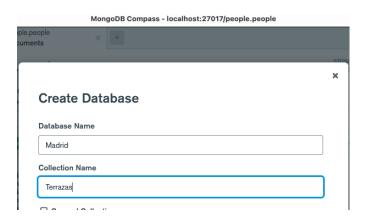
## 5. Caso de uso

5. Caso de uso: restricción de terrazas en Madrid por Covid-19.

Contratan a vuestro equipo para actualizar las restricciones de ciertos locales con terrazas en Madrid por cuestiones del Covid-19.

Los datos para actualizar están en el fichero llamado *act-grupal-openDataLocalesMadrid.cvs* y se os pide que consolidéis dichos cambios en una base de datos MongoDB llamada *Madrid* con la colección *Terrazas*. La información del dataset a utilizar lo tenéis en las siguientes URLs:

- Open Data Censo de locales, sus actividades y terrazas de hostelería y restauración (Terrazas)
- <a href="https://datos.madrid.es/FWProjects/egob/Catalogo/Economia/Ficheros/Estructura">https://datos.madrid.es/FWProjects/egob/Catalogo/Economia/Ficheros/Estructura</a> DS Fichero CLA Terrazas.pdf

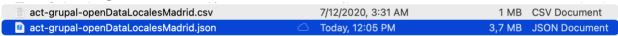


# 6. Conversion CSV a JSON

- 6.1. En el alcance del contrato se os pide:
  - a. Convertid el fichero CSV a formato JSON. Utilizad, por ejemplo, el servicio de https://csvjson.com/csv2json



- b. Descargad el fichero generado en el paso anterior, en formato JSON indicando la opción Array (todos los documentos estarán dentro de un único *array*).
  - i. En el fichero hay un problema de espacios en muchos campos String, intentad corregirlo en esta etapa.



Se identifican valores asociados a columnas sin valor, que se ubican a la decrecha de la utila columna "SILLAS\_RA", se proceden a eliminar dichas columnas.

2:S	sillas_ra 🔻	Column1 🔻	Column2 🔻	Column3 🔻	Column4 🔻	Column5 🔻	Column6 🔻	Column7 🔻	Column8 🔻
26									
48									
24	27:13:20	10:00:00	0:00:00	0	0	16	16	64	
42	F								

Se identifican espacios en blanco en columnas de tipo String, por lo cual se aplica la función rtrim() y ltrimp() a todas las columnas de tipo texto.

```
32;270068967;4;SALAMANCA SPACE
                                        ;RECOLETOS
                                                                 ; CALLE
33:270403150:4:SALAMANCA
                                        ; GUINDALERA
                                                                 ; AVENIDA
34;270078567;4;SALAMANCA
                                        ; GUINDALERA
                                                                 ; CALLE
36;40002021;4;SALAMANCA
                                       ; GUINDALERA
                                                                ; CALLE
         1 """
 In [ ]:
          2 Eliminar espacios en blanco a la derecha e izquierda de las columnas tipo object (string)
          3
          5 df_obj = df.select_dtypes(['object'])
          7 df[df_obj.columns] = df_obj.apply(lambda x: x.str.strip())
         1 """
 In [ ]:
          2 Eliminar columnas (_
                                1- 2 3 4 5 6 7) generadas dinamicamente por
          3 valores errados, hallados en el registro _id = 2812.
          6 df.drop(['__1', '__2', '__3','__4','__5','__6','__7'], axis=1, inplace=True)
 In [ ]: 1 df.tail(5)
          1 path = '/Users/juandavidescobarescobar/Documents/Unir/Materias/BD Big Data/Actividad 3/visa
 In [ ]:
            file = 'act-grupal-openDataLocalesMadrid-[corregido].csv'
          3
```

Se identifican campos numéricos con valores Null y vacíos, los cuales son remplazados por cero y cateados a tipo int.

4 df.to\_csv(path + file, index=0)

Adicional se identificaron valores atípicos en la variable numérica "**sillas\_ra**", por lo cual se remplazan los valores atípicos, nulos y vacíos por el valor o.

Se detallan todas las operaciones de limpieza y análisis de datos adicionales en el siguiente enlace: https://github.com/juadaves91/bd-bigdata-unir/blob/main/Activdad%203/Actividad3.ipynb

Resultado CSV corregido

```
pal-countries-small son act-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openDataLocalesMadrid_lact-grupal-openD
```

c. Editad el fichero generado y al principio del todo añadid la línea:

```
var datos_insertar = [...]
```

```
al-countries-small.js... | act-grupal-openDataLocales... | act
```

d. Guardad el fichero en una ruta conocida y cambiad su extensión a formato .js. ¿Sabéis qué tipo de fichero es un?j s? Java Script



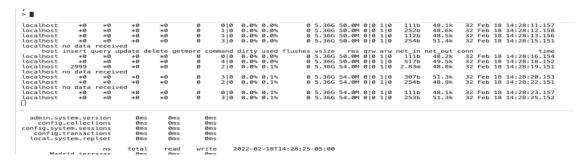
- e. Abrid el cliente mongo y ejecutad las siguientes instrucciones:
  - i. load("PATH\fichero.js")
  - ii. datos\_insertar[o]

f. Describid brevemente qué hacen las instrucciones anteriores.

**Load()** = Carga y ejecuta un archivo de tipo JavaScript en el cliente de mongo.

**datos\_insertar[o]** = Imprime por consola el primer documento almacenado en el array o colección definido en el archivo JavaScript.

### **Evidencia monitoreo:**



**Interpretación:** La operación mongostat —discover, reporta un log de estadísticas de todos los miembros replica set o sharded del cluster de MongoDB. La ventana de monitoreo registra en el log un total de 2999 registros insertados, 0.1% de uso. 48.8KB de uso de salida red, 2.83MB de entrada de red, tamaño de la operación 5.36 GB y la fecha de la transacción.

g. Con lo que habéis hecho antes, ¿podríais realizar un *insert* masivo?

Las instrucciones anteriores permiten cargar un array o colección de objetos tipo json o documentos, los cuales podría enviar como parámetro a al método insertMany(datos\_insertar)

h. Desde la terminal, podrías proponer una alternativa que evite realizar los pasos anteriores (descripción de dos líneas máximo y no vale indicar el uso MongoCompass (36)).

```
mongoimport --verbose --db madrid --collection terrazas --file
"/Users/juandavidescobarescobar/Documents/Unir/Materias/BD Big Data/Actividad
3/visa01 grupal/act-grupal-openDataLocalesMadrid-[corregido].json"
```

# 7. Consultas y actualizaciones en MongoDB

**6.2.** Las actualizaciones las llevaréis a cabo teniendo en cuenta bien la información del local en período estacional o la del resto del año. Las modificaciones son las siguientes:

- a. Los locales del barrio *Guindalera* de *Salamanca*, por motivos de la desescalada no podrán abrir y tendrán que permanecer cerrados. Cambiad la situación del local a *Cerrado* utilizando el *id* correspondiente.
  - Recordad que debéis constatar los tipos de situaciones que existan y seguir con la misma codificación.
  - A estos mismos locales, cambiad la situación de la terraza a cerrada siguiendo las mismas premisas anteriores.

```
Resultado consulta y monitoreo
                                                                                                                                                                                                                                                                                                                                               Consulta
                                                                                                                                                                                                                                                                                     db.terrazas.updateMany(
                                              Sset: {
                                                                                                                                                                                                                                                                                                                   desc_distrito_local:"SALAMANCA"
                                                desc_situacion_local :
id_situacion_local: 4
                                                                                                                                                                                                                                                                                                                   desc_barrio_local:"GUINDALERA"
                               : true, "matchedCount" : 32, "modifiedCount" : 32 }
                                                                                                                                                                                                                                 Preb 18 14:40:18.158
Preb 18 14:40:19.161
Preb 18 14:40:20.158
Preb 18 14:40:21.158
Preb 18 14:40:22.155
Preb 18 14:40:22.155
Preb 18 14:40:22.152
Preb 18 14:40:25.152
                                                                                                                                                                                                            48.3k
48.2k
49.1k
51.5k
49.4k
48.7k
48.6k
                                                                                                                                                                                                                            32
32
32
32
32
32
32
32
                                                                                                                                                                                                                                                                                                               id_situacion_terraza:8,
                                                                                                                                          res vsize res qrw arw
0 5.366 50.0M 0|0 1|0
                                                                                                                                                                                                                           tine
32 Feb 18 14:40:26.158
32 Feb 18 14:40:28.150
32 Feb 18 14:40:28.150
32 Feb 18 14:40:29.154
32 Feb 18 14:40:30.158
32 Feb 18 14:40:32.152
                                                                                                                                                                                             et_in
196b
253b
111b
112b
221b
167b
                                                                                                                                                                                                                                                                                                               desc_situacion_terraza: "Cerrado",
                                                                                                                                                                                                                                                                                                               desc_situacion_local : "Cerrado",
                                                                                                                                                                                                                                                                                                               id_situacion_local: 4
                                                                                                                                                                                                                                                                                           }
                                                                                       Ons
Ons
Ons
Ons
Ons
admin.system.version
config.collections
onfig.system.sessions
config.transactions
local.system.replset
                                                                                                       2022-02-18T14:40:30-05:00
```

- b. A todas las terrazas que se ubiquen en la *acera*, añadid un campo llamado *inspeccionar* y estableced el siguiente valor:
  - Si dispone de más de 10 mesas, true.
  - Si dispone de menos de 10 mesas, false.

```
Resultado consulta y monitoreo
                                                                                                                                                                                     Consulta
                                                                                                                               db.terrazas.aggregate([
                                                                                                                                       $addFields: {
                                                                                                                                                   total_mesas_es: { $add: [ "$mesas_es", "$mesas_aux_es"] },
                                                                                                                                                   total_mesas_ra: { $add: [ "$mesas_ra", "$mesas_aux_ra"] }
       l,
default: ""
                                                                                                                               )).forEach(function(x){
"matchedCount" : 2319, "modifiedCount" : 2319 }
                                                                                                                                                //print(x.total_mesas_es, x.total_mesas_ra)
                                                                                                                                                //print('--
                                                                             197b
111b
111b
222b
252b
253b
112b
111b
825b
                                                                                    50.8k
48.3k
48.2k
49.0k
48.6k
51.1k
48.4k
48.3k
49.1k
                                                                                                                                                  db.terrazas.save(x)
                                                                                                                               //update
                                                                                                                                db.terrazas.updateMany(
                                                                                                                                    $and: [
                                                                                                                                              desc_ubicacion_terraza:"Acera
                                                                                                                                }.
                              2022-02-18T14:50:11-05:00
                                                                                                                                  { $set: { inspeccionar: { $switch: {
                                                                                                                                                  branches: [
                                                                                                                                                      { case: { $gt: [ "total_mesas_es", 10 ] }, then: true },
                                                                                                                                                     { case: { $lte: [ "total_mesas_es", 10 ] }, then: false }, 
{ case: { $gt: [ "total_mesas_ra", 10 ] }, then: true },
                                                                                                                                                     { case: { $lte: [ "total_mesas_ra", 10 ] }, then: false }
                                                                                                                                                  default: ""
                                                                                                                                 }}}}
```

c. A las terrazas que se deban inspeccionar, asignad 2 mesas auxiliares y 8 sillas disponibles (*tanto en período estacional como el resto del año*).

```
Resultado consulta y monitoreo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Consulta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     db.terrazas.updateMany(
                                                                                              inspeccionar: true
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               inspeccionar: true
                                                                                                                                                                  mesas_aux_es: 2,
mesas_aux_ra: 2,
sillas_es: 8,
sillas_ra: 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               $inc: {
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             mesas_aux_es: 2,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             mesas_aux_ra: 2,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sillas_es: 8,
                                   ,
cknowledged" : true, "matchedCount" : 2319, "modifiedCount" : 2319 }
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sillas_ra: 8
localhost #0 #0
Localhost #0 #
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0 51.8k 32 Feb 18 14:52:57.102
48.5k 32 Feb 18 14:52:58.409
48.0k 32 Feb 18 14:52:59.106
49.4k 32 Feb 18 14:52:59.106
49.4k 32 Feb 18 14:53:00.103
48.5k 32 Feb 18 14:53:01.099
50.8k 32 Feb 18 14:53:03.103
48.8k 32 Feb 18 14:53:03.103
48.5k 32 Feb 18 14:53:05.103
net_out conn
48.3k 32 Feb 18 14:53:05.03
48.7k 32 Feb 18 14:53:06.103
                                                                                                                                                                                                            *0
*0
*0
*0
*0
*0
*0
*0
*0
*0
*0
*0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          602b
112b
111b
305b
112b
336b
112b
167b
166b
net_in
111b
112b
                                                                                                                                                              update
*0
*0
                                                                                                                                                                                                                                                                                                                                       mand
0|0
1|0
  localhost no data received localhost *0 *0 localhost *0 *0
  admin.system.version
config.collections
config.system.sessions
config.transactions
local.system.replset
                                                                                                                                                                                                                                               ens
ens
ens
ens
                                                                                                                                                                                                                                                                                                  Write
53ms
Oms
Oms
Oms
Oms
Oms
                                                                                                                                                                                  total
53ms
0ms
0ms
0ms
0ms
0ms
0ms
                                                                                                                                                                                                                                                                                                                                                           2022-02-18T14:53:11-05:00
                                                                                                                                                                                                                                             read
Oms
Oms
Oms
Oms
Oms
Oms
  ns
Madrid.terrazs
Madrid.zona1
admin.system.version
config.collections
config.system.sessions
config.transactions
local.system.replset
```

- d. A las terrazas que no deban ser inspeccionadas, añadid el campo estado con el valor:
  - 1 si el número de sillas es menor que 10.
  - 2 si el número de sillas está entre 10 y 20.
  - 3 si cuenta con más de 20 sillas.

```
Resultado consulta y monitoreo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Consulta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       //campos calculados
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       db.terrazas.aggregate([
                                                                                                                                                                                                                                                                                                                                                                                                                                                   { $lt : [ "total_sillas", 20 ] } ]
}, then: 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            total_sillas: { $add: [ "$sillas_es", "$sillas_ra"] }
                                                                                                                                                                                                                                                                                                                                                                                                                     case: { sqt: [ "total sillas", 20 ] }, then: 3
                                                                                                                                                                                                                                                                                                                              ],
default: 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          )
]).forEach(function(x){
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         print(x.total_sillas)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         //print('-----')
db.terrazas.save(x)
                                                                                       }
                                                                                                                             : true, "matchedCount" : 0, "modifiedCount" : 0 }
localhost no data localhost +80 localhost +8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0 5.386 63.0% 0 0 10 10 0 5.386 63.0% 10 10 10 0 5.386 63.0% 10 10 10 0 5.386 63.0% 10 10 10 0 5.386 63.0% 10 10 10 0 5.386 63.0% 10 10 10 0 5.386 63.0% 10 10 10 0 5.386 63.0% 10 10 10 0 5.386 63.0% 10 10 10 0 5.386 63.0% 10 10 10 0 5.386 63.0% 10 10 10 0 5.386 63.0% 10 10 10 0 5.386 63.0% 10 10 10 0 5.386 63.0% 10 10 10 0 5.386 63.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 10 0 5.386 64.0% 10 10 0 5.386 64.0% 10 10 0 5.386 64.0% 10 10 0 5.386 64.0% 10 10 0 5.086 64.0% 10 10 0 5.086 64.0% 10 10 0 5.386 64.0% 10 10 0 5.086 64.0% 10 0 5.086 64.0% 10 0 5.086 64.0% 10 0 5.086 64.0% 10 0 5.086 64.0% 10 0 5.086 64.0% 10 0 5.086 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       //update
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Feb 18 14:56:38.107
Feb 18 14:56:39.106
Feb 18 14:56:40.107
Feb 18 14:56:41.105
Feb 18 14:56:42.107
Feb 18 14:56:43.108
Feb 18 14:56:44.103
Feb 18 14:56:45.102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         48.9k
48.4k
48.9k
48.9k
48.9k
49.3k
48.7k
iet_out
48.4k
50.6k
48.5k
49.3k
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       db.terrazas.updateMany({
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             inspeccionar : false
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        107b
et_in
112b
197b
112b
514b
221b
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          32
32
32
32
32
32
32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      valor: {
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sswitch: {
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     case: {$It: ["total sillas", 10]}, then: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           case: { Sand : [{Sgte: ["total sillas", 10]}.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              {$lt:["total_sillas", 20]}]
                                                                                                                                                                                                 total
ens
ens
ens
ens
ens
ens
                                                                                                                                                                                                                                                                                                                        write
Oms
Oms
Oms
Oms
Oms
Oms
                                                                                                                                                                                                                                                                                                                                                                                          2022-02-18T14:56:51-05:00
                                                                                                                                                                                                                                                               read
Ons
Ons
Ons
Ons
Ons
Ons
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        case: { $gt: [ "total_sillas", 20 ] }, then: 3
```

e. De lunes a jueves, ningún local podrá cerrar más allá de las 00:00:00, actualizad el horario de cierre a esta nueva hora límite.

Resultado consulta y monitoreo	Consulta
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL  > db.terrazas.updateMany{{},{	db.terrazas.updateMany({},{ \$set: { "hora_fin_U_es": "00:00:00", "hora_ini_U_ra": "00:00:00" } })
localhost	
admin.system.version @ms @ms @ms @ms @ms config.collections @ms @ms @ms @ms @ms config.stem.sessions @ms @ms @ms @ms config.stransactions @ms @ms @ms @ms @ms (acal.system.replacet @ms @ms @ms @ms @ms (acal.system.replacet @ms	

f. De viernes a sábado, los locales que cierren a las 2:30:00, ahora tendrán que hacerlo a las 2:00:00.

Resultado consulta y monitoreo	Consulta
> db.terrazas.updateMany{{	db.terrazas.updateMany({     hora_fin_VS_es: "02:30:00" },{     \$set: { "hora_fin_VS_es" : "02:00:00"} }}) db.terrazas.updateMany({     hora_fin_VS_ra: "02:30:00"
localhost	}, {     \$set: { hora_fin_VS_ra : "02:00:00"} })
admin.system.version on ons ons ons ons ons ons ons ons on	

g. A todos los locales que estén sobre la calle Alcalá en Madrid se les debe inspeccionar.

Resultado consulta y monitoreo	Consulta
•	db.terrazas.updateMany(
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL	1
db.terrazas.updateMany(	L
	\$and: [
. Şand: [	1
. "DESC_CLASE": "CALLE"	
. },	"DESC_CLASE": "CALLE"
. {	},
. "DESC_NOMBRE": "ALCALA"	, ", ", ", ", ", ", ", ", ", ", ", ", ",
. 1	1
	"DESC_NOMBRE":"ALCALA
. 3/	ı -
. { Sset: { "inspeccionar": true } }); "acknowledged" : true, "matchedCount" : 50, "modifiedCount" : 1 }	I I
"acknowledged" : true, "matchedLount" : 50, "modifiedCount" : 1 }	]
-	3.
calhost *0 *0 *0 *0 *0 0 0 0 0.0% 0.1% 0 5.35G 40.0M 0 0 1 1 1 1 1 1 1 48.3	
calhost *0 *0 *0 *0 *0 0 1 0 0.0% 0.1% 0 5.35G 40.0M 0 0 1 0 112b 48.5 calhost *0 *0 *0 *0 0 2 0 0.0% 0.1% 0 5.35G 40.0M 0 0 1 0 251b 50.9	
calhost *8 *8 *8 *8 *8 8 2 2 8 8 8 1 1 5 1.35 40.00 1 1 1 1 1 1 1 1 5 1 5 1 5 1 5 1 5 1 5	
host insert query update delete getmore command dirty used flushes vsize res grw arw net in net ou	t conn time
calhost ×0 ×0 ×0 ×0 0 2 0 0.0% 0.1% 0 5.35G 40.0M 0 0 1 0 253b 48.7	k 32 Feb 18 16:56:46.889
calhost no data received calhost *8 *8 *8 *8 0 2 0 0.0% 0.1% 0 5.35G 40.0M 0 0 1 0 199b 51.1	k 32 Feb 18 16:56:48.889
calnost %8 %8 %90 %8 8 2 2 8 8 8 5.1.15 8 5.1.25 40.04 8 8 19 51.1.	X 32 FED 18 16:50:46.889
calhost ×0 ×0 ×0 ×0 0 2 0 0.0% 0.1% 0 5.35G 40.0M 0 0 1 0 220b 48.7	k 32 Feb 18 16:56:49.894
calhost ×0 ×0 ×0 ×0 0 2 0 0.0% 0.1% 0 5.35G 40.0M 0 0 1 0 253b 48.8	
calhost ×0 ×0 ×0 ×0 0 1 0 0.1% 0.1% 0 5.35G 40.0M 0 0 1 0 454b 48.4	
calhost *0 *0 *0 *0 0 1 0 0.1% 0.1% 0.5.35G 40.0M 0 0 1 0 112b 48.6 calhost *0 *0 *0 *0 0 2 0 0.1% 0.1% 0.5.35G 40.0M 0 0 1 0 252b 51.0	
calhost *0 *0 *0 *0 0 2 0 0.1% 0 5.35G 40.0M 0 0 1 0 157b 48.7	
admin.system.version @ms @ms @ms	
config.collections @ms @ms @ms	
nfig.system.sessions @ms @ms @ms	
config.transactions 0ms 0ms 0ms local.system.replact 0ms 0ms 0ms	
tocat.system.reptset ons ons	
ns total read write 2022-02-18T16:56:53-05:00	
Madrid.terrazas 8ms 0ms 8ms	
Madrid zonal Ome Ome Ome	
	<b> </b>

h. A todos los locales con terrazas abiertas, añadid un campo *revision* cuyo valor sea un documento con la siguiente información: {prox\_inspeccion: 10, puntuacion: 80, comentario: "separar las mesas"}

Resultado consulta y monitoreo	Consulta
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL  > db.terrazas.updateMany(	db.terrazas.updateMany( {
Coralhost   *0	puntuacion: 80, comentario: "separar las mesas' } }};
Badish.system.version	

i. Cread una nueva colección llamada *zona1*, con todos los locales que pertenezcan al distrito de Villaverde.

				F	Resultado cons	ulta y monitored	)					Consulta
PROBLEMS	OUTPUT	DEBUG CONS	OLE	TERMINAL							// Creat collection	n
	2561, 2599, 2600, 2771, 2779, 3105,										var array_zona1 : \$and	= db.terrazas.find({ :[ {
	3124, 3365, 3608, 3750,											id_distrito_local : 17 },
	3752, 3757, 3881											{ desc distrito local
											"VILLAVERDE"	,
ocalhost	*0 ×0 *0 *0 *0 *0 *0 *0	*8 *8 *8 *8 *8 we update dele *8 *8 *8 ived *8 *8 *8	*8 *8 *8 *8 *8 *8 *8 *8 *8 *8 *8 *8 *8 *	0 0 1 0 0 nore comm	2   8	8 5.366 45.00 8 8 8 5.366 45.00 18 8 5.366 45.00 18 8 5.366 45.00 18 8 5.366 45.00 18 10 5.396 45.00 18 10 5.396 45.00 18 10 5.396 45.00 18 10 5.396 45.00 18 10 5.396 45.00 18 10 5.396 45.00 18 10 5.396 46.00 1	1   0   166b   1   0   253b   1   0   111b   1   0   111b   1   0   167b   1   0   111b   1   0   111b   1   0   113b   1   0   157b   1   0   115b   1   0   1   0   115b   1   0   1   0   115b   1   0   0	48.9k 48.4k 48.8k 48.2k 184k 48.9k net_out c 48.3k 48.3k 49.2k 49.2k 48.7k 48.1k 48.1k 48.9k 48.4k	32 Feb 18 16 32 Feb 18 17 32 Feb 18 17 32 Feb 18 17 32 Feb 18 17 32 Feb 18 17	:59:50.898 :59:51.895 :59:52.898 :59:53.897 :59:54.893 time :59:55.893 :59:55.893 :59:55.893 :59:55.897 :59:59.892 :00:80.897		
onfig.syste	collections m.sessions ansactions	ens ens ens ens	Ons Ons Ons Ons	Oms Oms Oms Oms								
	ns id.terrazas idrid.zonal en.version	total Øns Øns	read Oms Oms	write Oms Oms	2022-02-18T17:0	08:03-05:00						

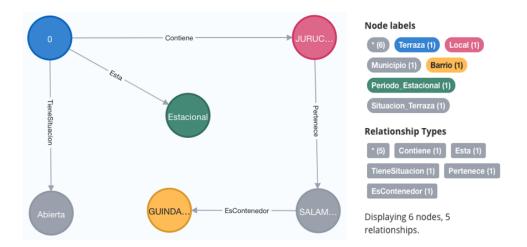
j. Cread una nueva colección llamada *zona2*, con todos los locales que pertenezcan al distrito de Salamanca y barrio Castellana.

Las restricciones se aplican sin importar la estación del año.

						Resu	ltado	con	sulta y mo	onito	oreo								Consulta
																			var array zona2 = db.terrazas.find({
PROBLEMS	OUTP	UT I	DEBUG COL	NSOLE	TERMI	NAL													\$and : [
	36	61,				_													şanu . [
	36	88,																	
		92, 93,																	{"desc distrito local" : "SALAMANC
	31	100.																	
		106,																	
		118,																	{"id_distrito_local" : 4},
		58,																	
		26,																	{"desc barrio local": "CASTELLANA"
		54, 18.																	desc_barro_local : CASTELLANA
		56																	
1																			1
																			}, {}).toArray()
•																			), ()).toArray()
ocalhost n				_	_														
ocalhost ocalhost	*0	*0 *0	*0 *0	*0 *0	9		0.0%				.0M 0   0 .0M 0   0		111b 335b	48.1k 51.5k			8 17:01:02.8 8 17:01:03.8		db.zona2.insertMany(array_zona2)
ocalhost	+6	1	*0	+0	0	3 0	0.0%	0.1%	0 5.3	5G 46	.am e e		511b	97.7k			8 17:01:04.8		,, ,_ ,_ ,_ ,_
host i			pdate de			command	dirty	used	flushes vsi	ze	res qrv	arw r	net_in	net_out				ime	
ocalhost ocalhost n	*8	*0	*3	*8	0	1 0	0.09	0.1%	0 5.3	5G 46	.am e e	1 0	112b	48.4k	32	Feb 18	8 17:01:06.8	893	// Validar resultados
ocalhost	*6	*9	*9	*0	9	410	0.0%	0.1%	0 5.3	5G 46	.am e   e	1 0	568b	49.4k	32	Feb 1	8 17:01:07.8	896	
ocalhost	*0	*9	*0	*0	0		0.0%				. ам е је		197b	50.7k	32	Feb 1	8 17:01:08.8	898	db.zona2.find({}).pretty()
ocalhost ocalhost	+0 43	*0	*8	+0	9		0.0%				.am e je .am e je		167b 48.6k	48.7k 48.6k			8 17:01:09.8 8 17:01:10.8		db.zonaz.find({}).pretty()
ocalhost	*8	*8	*3	+0	0	2   6	0.03	0.1%	0 5.3	5G 46	. am e   e	1 0	253b	48.7k	32	Feb 18	8 17:01:11.8	898	
ocalhost	*0	*9	*9	*0	0		0.0%				. am e je		112b	48.5k			8 17:01:12.8		
ocalhost ocalhost	*8	*8 *8	*0 *0	*0	9		0.0%				.aM 0   6 .aM 0   6		252b 167b	51.4k 48.7k			8 17:01:13.8 8 17:01:14.8		
oca tilos c	70	-70				210	2101	0.10	0 313	40		-10	2370	70171	32		2,.02,1410		
admin.sys	tem.ver	sion	Øms	ens.		ens													
config.	collect	ions	0ms	8ms	, (	Ins													
onfig.system config.t			Ons Ons	ens ens		ens ens													
local.sys			Ons	ens ens		ons Ons													
,-																			
onfig.syst	on. corr	ns	total 1ms	read		lte Ins	2022-0	2-18T1	7:01:13-05:	99									
	em.sess id.terr		Ons	ens		uns Ans													
	adrid.z		0ms	ens		ens													

# 8. Neo4j

- 6. En Neo4J, proponed un modelo de datos donde sea posible visualizar los locales de cada barrio y la tipología de terrazas que tienen. Indicad los nodos y sus relaciones; una vez modelado, cread una única consulta que visualice el grafo total. Añadid tanto a nodos como a relaciones sus atributos, para ello elegid los campos que más convengan a vuestro modelo. **De este punto se espera**:
  - a. Un grafo genérico que represente vuestro modelo (nodos, relaciones y atributos; se deben reflejar todos los campos utilizados).



b. Una breve descripción del modelo propuesto.

Modelo entidad relación generado a partir de entidades identificadas de la sabana de datos, act-grupalopenDataLocalesMadrid.xlsx. Las entidades que se generan son las siguientes:

- Terrazas
- Locales
- Periodo estacional
- Situación terrazas
- Municipios
- Barrios

Dichos nodos o entidades se asociaron mediante las siguientes relaciones:

- **Terraza Local:** Contiene
- Local Municipio: Pertenece
- Municipio Barrio: EsContenedor
- Terraza Periodo\_Estacional: Esta
- Terraza Situacion\_Terraza: TieneSituacion

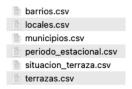
c. La descripción del proceso utilizado para cargar los datos en Neo4J a partir de los datos en MongoDB.

**Paso 1.** Lo primero que se hizo fue un proceso de normalización de los datos, identificando primero las entidades que se encuentran en la sabana de datos act-grupal-openDataLocalesMadrid.xlsx, la siguiente tabla muestra la agrupación de cada uno de los 36 campos en las entidades o tablas identificadas:



Cada una de las tablas tiene un identificador excepto la tabla Barrio, la cual contiene únicamente el nombre o descripción del barrio, estandarizado previamente, mediante un proceso en Jupyter Nootbok.

**Paso 2.** El siguiente paso fue atreves de Python y la librería de Pandas generar, los conjuntos de datos de cada una de las entidades, y se exportan en archivos con formato csv, los cuales se listan a continuación, para ver mas detalle de la operación mencionada se puede referir al adjunto **Actividad 3. Ipny.** 



**Paso 3.** A continuación, se crean conjuntos de datos en Python los cuales soportan o almacenan los datos de las llaves que generan la relación entre 2 nodos, a continuación, se listan las tablas creadas.



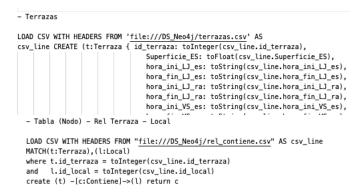
Los conjuntos de datos generados en el **Paso 2 y 3**, fueron exportados en formato CSV, en la ruta configurada por defecto para la BD de NEO4J creada para el ejercicio



Paso 3. En Neo4j se crea un nuevo proyecto "Actividad 3" y un nueva Base de datos llamada "Censo M"



**Paso 4.** En **Neo4j** Se generan las sentencias de creación de Nodos y relaciones, las cuales se encuentran descritas en el archivo anexo **ESCOBAR\_ESCOBAR\_JUAN\_DAVID\_actividad\_3\_(Grupal).txt** 



**Paso 4.** Al momento de ejecutar las sentencias de creación de relaciones entre nodos, se puede generar un error de desbordamiento de memoria, por lo cual se opta por aumentar en tamaño de disponibilidad de memoria para Neo4j en el archivo de configuración, o trabajar con menos registros.

d. Una consulta CQL y capturas con la visualización del grafo generado con los datos.

```
MATCH (t:Terraza)-[Contiene]->(l:Local),

(l:Local)-[Pertenece]->(m:Municipio),

(m:Municipio)-[EsContenedor]->(b:Barrio),

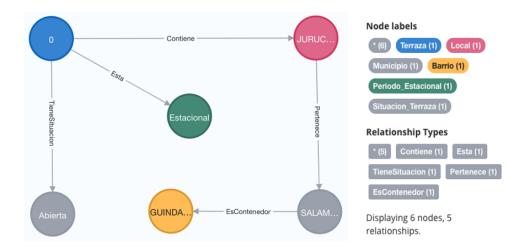
(t:Terraza)-[Esta]->(pe:Periodo_Estacional),

(t:Terraza)-[TieneSituacion]->(st:Situacion_Terraza)

WHERE t.id_terraza = 32

RETURN t, l, m, b, pe, st

LIMIT 1
```



## 9. Anexos

Script ETL: https://github.com/juadaves91/bd-bigdata-unir/blob/main/Activdad%203/Actividad%203.ipynb

Script MongoDB: <a href="https://github.com/juadaves91/bd-bigdata-unir/blob/main/Activdad%203/ESCOBAR ESCOBAR JUAN DAVID Actividad Grupal.txt">https://github.com/juadaves91/bd-bigdata-unir/blob/main/Activdad%203/ESCOBAR ESCOBAR JUAN DAVID Actividad Grupal.txt</a>

Script Neo4j: https://github.com/juadaves91/bd-bigdata-unir/blob/main/Activdad%203/ESCOBAR ESCOBAR JUAN DAVID Actividad Grupal.txt

 $\begin{tabular}{ll} \textbf{Modelo entidad relación:} & $\underline{\text{https://github.com/juadaves91/bd-bigdata-unir/blob/main/Activdad%203/Actividad%203.ipynb} \\ \end{tabular}$