



DEPARTAMENTO
DE COMPUTACION

Facultad de Ciencias Exactas y Naturales - UBA

Trabajo Práctico #2

16 de Septiembre de 2016

Base de Datos

Integrante	LU	Correo electrónico
Pedro Rodriguez	197/12	pedro3110.jim@gmail.com
Lucas Tavoraro Ortiz	322/12	tavo92@gmail.com



Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires

Ciudad Universitaria - Pabellón I

Intendente Güiraldes 2160 - C1428EGA

Ciudad Autónoma de Buenos Aires - Argentina

Tel/Fax: (54 11) 4576-3359

<http://exactas.uba.ar>

Índice

1. Introducción	3
2. Documentos para Mongo	3
2.1. Querie 1 + 2	3
2.2. Querie 3	3
2.3. Querie 4	3
2.4. Querie 5+6	4
2.5. Querie 7	4
3. Conclusiones	4

1. Introducción

2. Documentos para Mongo

A continuación vamos a definir el documento para Mongo en cada una de las queries dadas y argumentar la elección de dicho documento contando, a alto nivel, como serían los inserts, mantenimiento y búsqueda con map-reduce.

2.1. Querie 1 + 2

1. Para mayor nro. de crímenes cometidos por alguna persona 2. Para nro de crímenes promedio cometidos por criminales

Para estas primeras dos queries, elegimos diseñar el documento expuesto a continuación: Pensamos que lo más importante es optimizar las búsquedas (ie. los map-reduce). Para esto, tomamos como clave de cada posible documento el **dniCulpable** de la persona que fue encontrada culpable de una cantidad **cantCasosCulpable** crímenes.

Inserts: Cada vez que una persona es hallada culpable de un nuevo caso, hay que acceder al documento que le corresponde a ese dni a través de la key **dniCulpable**, e incrementar su valor **cantCasosCulpable**.

Mantenimiento: cada documento ocupa muy poco espacio, con lo cuál es fácilmente mantenible y escalable.

Map Reduce: Para las queries 1 y 2 el map mapearía todos los valores a un mismo key, y el reduce buscaría el máximo y el promedio de dichos valores respectivamente.

```
1      [
2          {
3              dniCulpable: 37206752,
4              cantCasosCulpable: 2
5          }
6      ]
```

2.2. Querie 3

3. Para 10 ciudades con mayor nro. de crímenes Elegimos el diseño expuesto a continuación: para cada ciudad, guardamos su id **idCiudad** como clave, y como valor una lista **dniCulpables** con los casos c

```
1      [
2          {
3              idCiudad: 4,
4              dniCulpables:
5                  [
6                      {dniCulpable: 37206752},
7                      {dniCulpable: 35222345}
8                  ]
9          }
10     ]
```

2.3. Querie 4

Cuántas personas se han visto involucradas como testigos en el mayor número de casos

```
1      [
2          {
3              dniTestigo: 27893435,
4              idCasosInvolucradaComoTestigo:
5                  [
6                      {idCaso: 1},
7                      {idCaso: 4},
```

```

8                                     {idCaso: 6}
9                                     ]
10                                }
11    ]

```

2.4. Querie 5+6

5. Para casos con mayor nro. de involucrados 6. Para casos con mayor nro. de evidencia

```

1    [
2        {
3            idCaso: 4,
4            {
5                idEvidencias:
6                    [
7                        {idEvidencia: 3},
8                        {idEvidencia: 4}
9                    ],
10               dniInvolucrados:
11                   [
12                       {dniInvolucrado: 37206752}
13                   ]
14            }
15        }
16    ]

```

2.5. Querie 7

7. Para cantidad de crímenes de cada tipo cometidos en los últimos 45 días Para optimizar el uso del disco, creamos un documento para cada tipo de crimen. Cada documento tiene como valor un diccionario que para cada fecha de los últimos 45 días, tiene como valor la cantidad de crímenes que se cometieron ese día. Mantenimiento: Cada día, hay que recorrer el diccionario correspondiente a cada tipo de crimen, y borrar las entradas correspondientes a fechas de hace más de 45 días. La ventaja de esto, es que sabemos cuánta memoria ocupa cada documento, pues como máximo puede tener 45 fechas. Inserts: Buscar en el tipo de crimen correspondiente, buscar en el diccionario la key correspondiente a la fecha, e incrementar el valor. Map Reduce: El map recorre el diccionario correspondiente al tipo de crimen a buscar, y mapea sobre una misma clave, la suma de todos los crímenes en las fechas que hay guardadas (todas están incluidas en los últimos 45 días).

```

1    [
2        {
3            idTipoCrimen: 4,
4            cantCrimenesPorFecha:
5                {
6                    {fecha: 2014-10-08,
7                      cantidadDeCrimenes: 4},
8                    {fecha: 2015-03-10,
9                      cantidadDeCrimenes: 1},
10                   {fecha: 2016-10-10,
11                     cantidadDeCrimenes: 2}
12                }
13        }
14    ]

```

3. Conclusiones