Módulo 3: Bases de Datos NO SQL



Alumno: Pablo Pérez Calvo

1. Analizar con find la colección.

```
1 // 1. db.movies.find()
```

```
movies 🕝 0.421 s 🤪 Fetch Count
                                                             ₹ -
1 /* 1 createdAt:16/12/2024 14:03:05*/
  2 - {
         "_id" : ObjectId("676025098307048abbe4772f"),
 3
         "title" : "Caught",
 4
         "year" : 1900,
         "cast" : [ "Undefined" ],
  6
          "genres" : [ "Undefined" ]
 8 },
 9
 10 /* 2 createdAt:16/12/2024 14:03:05*/
11 - {
         "_id" : ObjectId("676025098307048abbe47730"),
"title" : "After Dark in Central Park",
 12
 13
         "year" : 1900,
"cast" : [ "Undefined" ],
 14
 15
 16
         "genres" : [ "Undefined" ]
 17 },
 18
 19 /* 3 createdAt:16/12/2024 14:03:05*/
 20 - {
 21
         "_id" : ObjectId("676025098307048abbe47731"),
         "title" : "Buffalo Bill's Wild West Parad",
22
         "year" : 1900,
 23
         "cast" : [ "Undefined" ],
"genres" : [ "Undefined" ]
 24
 25
 26 },
27
```

Figura 1: Resultado 1

2. Contar cuántos documentos (películas) tiene cargado.

```
// 2.
db.movies.countDocuments()
```



Figura 2: Resultado 2

3. Insertar una película.

```
// 3.
db.movies.insertOne({title: "Once pares de botas", year: 1954, cast : [], genres: []})
```

Figura 3: Resultado 3

4. Borrar la película insertada en el punto anterior (en el 3).

```
1 // 4.
2 db.movies.deleteOne({title: "Once pares de botas"})
```

Figura 4: Resultado 4

5. Contar cuantas películas tienen actores (cast) que se llaman "and". Estos nombres de actores están por ERROR.

```
// 5.
db.movies.find({cast: "and"}).count()
```



Figura 5: Resultado 5

6. Actualizar los documentos cuyo actor (cast) tenga por error el valor "and" como si realmente fuera un actor. Para ello, se debe sacar únicamente ese valor del array cast. Por lo tanto, no se debe eliminar ni el documento (película) ni su array cast con el resto de actores.

```
// 6. db.movies.updateMany({cast: "and"}, { $pull: { cast: "and" } })
```

```
1 - {
2          "acknowledged" : true,
3          "matchedCount" : 93,
4          "modifiedCount" : 93
5 }
```

Figura 6: Resultado 6

7. Contar cuantos documentos (películas) tienen el array 'cast' vacío.

```
// 7.
db.movies.countDocuments({cast: {$size: 0 }})
```



Figura 7: Resultado 7

8. Actualizar TODOS los documentos (películas) que tengan el array cast vacío, añadiendo un nuevo elemento dentro del array con valor Undefined. Cuidado! El tipo de cast debe seguir siendo un array. El array debe ser así -> ["Undefined"].

```
// 8.
db.movies.updateMany({cast: {$size: 0 }}, { $set: {cast: ["Undefined"]}})
```

```
0.591 s

1 - {

2          "acknowledged" : true,

3          "matchedCount" : 986,

4          "modifiedCount" : 986

5 }
```

Figura 8: Resultado 8

9. Contar cuantos documentos (películas) tienen el array genres vacío.

```
// 9.
db.movies.countDocuments({genres: {$size: 0 }})
```



Figura 9: Resultado 9

10. Actualizar TODOS los documentos (películas) que tengan el array genres vacío, añadiendo un nuevo elemento dentro del array con valor Undefined. Cuidado! El tipo de genres debe seguir siendo un array. El array debe ser así -> ["Undefined"].

```
// 10.
db.movies.updateMany({genres: {\$size: 0 }}, {\$set: {genres: ["Undefined"]}})
```

Figura 10: Resultado 10

11. Mostrar el año más reciente / actual que tenemos sobre todas las películas.

```
// 11.
db.movies.find({}, { _id: 0, year: 1 }).sort({ year: -1 }).limit(1)
```



Figura 11: Resultado 11

12. Contar cuántas películas han salido en los últimos 20 años. Debe hacerse desde el último año que se tienen registradas películas en la colección, mostrando el resultado total de esos años. Se debe hacer con el Framework de Agregación.

Figura 12: Resultado 12

13. Contar cuántas películas han salido en la década de los 60 (del 60 al 69 incluidos). Se debe hacer con el Framework de Agregación.

```
// 13.
var query1 = {year: {$gte: 1960, $lte:1969}}
var fase1 = {$match: query1}
var query2 = {_id: null, title: {$sum: 1}}
var fase2 = {$group: query2}
var etapas = [fase1, fase2]
db.movies.aggregate( etapas)
```

Figura 13: Resultado 13

14. Mostrar el año u años con más películas mostrando el número de películas de ese año. Revisar si varios años pueden compartir tener el mayor número de películas.

```
// 14.
var query1 = {_id: "$year", pelis: {$sum: 1}}
var fase1 = {$group: query1}
var query2 = {"pelis":-1}
var fase2 = {$sort: query2}
var fase3 = {$limit:1}
var etapas = [fase1, fase2, fase3]
db.movies.aggregate( etapas)
```

Figura 14: Resultado 14

15. Mostrar el año u años con menos películas mostrando el número de películas de ese año. Revisar si varios años pueden compartir tener el menor número de películas.

```
// 15.
var query1 = {_id: "$year", pelis: {$sum: 1}}
var fase1 = {$group: query1}
var query2 = {"pelis":1}
var fase2 = {$sort: query2}
var fase3 = {$limit:3}
var etapas = [fase1, fase2, fase3]
db.movies.aggregate(etapas)
```

```
movies | 0.441 s | 3 Docs
1 /* 1 */
  2 ₹ {
        "_id" : 1906,
  3
        "pelis" : 7
  4
  5 },
  6
 7 /* 2 */
 8 + {
 9
        "_id" : 1907,
        "pelis" : 7
 10
 11 },
 12
 13 /* 3 */
 14 ▼ {
        "_id" : 1902,
 15
        "pelis" : 7
 17 }
```

Figura 15: Resultado 15

16. Guardar en nueva colección llamada "actors" realizando la fase \$unwind por actor. Después, contar cuantos documentos existen en la nueva colección.

```
// 16.
var fase1 = {$unwind: "$cast"}
var query2 = {"_id": 0}
var fase2 = {$project: query2}
var fase3 = { $out: "actors"}
var etapas = [fase1,fase2, fase3]
db.movies.aggregate(etapas)

db.actors.countDocuments()
```



Figura 16: Resultado 16

17. Sobre actors (nueva colección), mostrar la lista con los 5 actores que han participado en más películas mostrando el número de películas en las que ha participado. Importante! Se necesita previamente filtrar para descartar aquellos actores llamados "Undefined". Aclarar que no se eliminan de la colección, sólo que filtramos para que no aparezcan.

```
// 17.

var query1 = {cast: {$ne: "Undefined"} }

var fase1 = {$match: query1}

var query2 = { _id: "$cast", cuenta: {$sum: 1}}

var fase2 = {$group: query2}

var query3 = {cuenta: -1}

var fase3 = {$sort: query3}

var fase4 = {$limit: 5}

etapas = [fase1, fase2, fase3, fase4]

db.actors.aggregate(etapas)
```

Figura 17: Resultado 17

18. Sobre actors (nueva colección), agrupar por película y año mostrando las 5 en las que más actores hayan participado, mostrando el número total de actores.

```
// 18.
var query1 = { _id: {title:"$title", year: "$year"}, cuenta: {$sum : 1}}
var fase1 = {$group: query1}
var query2 = {cuenta: -1}
var fase2 = {$sort: query2}
var fase3 = {$limit: 5}
etapas = [fase1, fase2, fase3]
db.actors.aggregate(etapas)
```

```
"_id" : {
| "title" : "The Twilight Saga: Breaking Dawn - Part 2",
| "year" : 2012
                                                                  28 /* 4 */
      },
"cuenta" : 35
                                                                  29 + {
30 + "_id" : {
                                                                          "title" : "Avengers: Infinity War",

"year" : 2018
                                                                  31
32
     "cuenta" : 29
                                                                  35 },
                                                                  38 + {
                                                                         "_id" : {
    "title" : "Grown Ups 2",
    "year" : 2013
},
                                                                  39 ÷
    "cuenta" : 28
                                                                  44 }
```

Figura 18: Resultado 18

19. Sobre actors (nueva colección), mostrar los 5 actores cuya carrera haya sido la más larga. Para ello, se debe mostrar cuándo comenzó su carrera, cuándo finalizó y cuántos años ha trabajado. Importante! Se necesita previamente filtrar para descartar aquellos actores llamados "Undefined". Aclarar que no se eliminan de la colección, sólo que filtramos para que no aparezcan.

```
// 19.
   var query1 = {cast: {$ne: "Undefined"} }
2
   var fase1 = {$match: query1}
   var query2 = { _id: "$cast", comienza: {$min: "$year"}, termina: {
4
       $max: "$year"}}
   var fase2 = {$group: query2}
5
   var query3 = {anos: {$add:[ { $subtract: ["$termina", "$comienza
6
       "]},1]}}
   var fase3 = {$addFields: query3}
   var query4 = {anos: -1}
   var fase4 = {$sort: query4}
9
   var fase5 = {$limit: 5}
10
11
   etapas = [fase1, fase2, fase3, fase4, fase5]
   db.actors.aggregate(etapas)
```

```
"_id" : "Harrison Ford",
                                               /* 4 */
   "comienza" : 1919.
   "termina" : 2017,
   "anos" : 99
                                               · - {
                                                       "_id" : "Lillian Gish",
                                                       "comienza" : 1912,
/* 2 */
                                                       "termina" : 1987,
   "_id" : "Gloria Stuart",
                                                       "anos" : 76
   "comienza" : 1932,
                                                 },
   "termina" : 2012,
   "anos" : 81
                                               · + {
                                                       "_id" : "Angela Lansbury",
   "_id" : "Kenny Baker",
                                                       "comienza" : 1944,
   "comienza" : 1937,
                                                       "termina" : 2018,
   "termina" : 2012,
                                                       "anos" : 75
   "anos" : 76
```

Figura 19: Resultado 19

20. Sobre actors (nueva colección), Guardar en nueva colección llamada "genres" realizando la fase \$unwind por genres. Después, contar cuantos documentos existen en la nueva colección.

```
// 20.
var fase1 = {$unwind: "$genres"}
var query2 = {"_id": 0}
var fase2 = {$project: query2}
var fase3 = { $out: "genres"}
var etapas = [fase1,fase2, fase3]
db.actors.aggregate(etapas)

db.genres.countDocuments()
```



Figura 20: Resultado 20

21. Sobre genres (nueva colección), mostrar los 5 documentos agrupados por "Año y Género" que más número de películas diferentes tienen mostrando el número total de películas.

```
1 /* 1 */
        " id" : {
3 →
            "year" : 1919,
4
            "genre" : "Drama"
5
6
        "pelis" : 291
7
8 },
                                                 28 /* 4 */
                                                 29 - {
11 - {
                                                 30 +
                                                         "_id" : {
        "_id" : {
12 -
                                                 31
                                                             "year" : 1919,
            "year" : 1925,
13
                                                             "genre" : "Comedy"
           "genre" : "Drama"
                                                 32
14
                                                 33
15
                                                         "pelis" : 226
        "pelis" : 247
                                                 34
16
                                                35 },
17 },
                                                 36
18
                                                 37 /* 5 */
19 /* 3 */
                                                 38 - {
20 - {
       "_id" : {
    "year" : 1924,
    "genre" : "Drama"
                                                 39 +
                                                         "_id" : {
21 -
                                                             "year" : 1922,
                                                 40
22
                                                 41
                                                            "genre" : "Drama"
23
                                                 42
                                                         "pelis" : 209
                                                 43
        "pelis" : 233
                                                 44 }
```

Figura 21: Resultado 21

22. Sobre genres (nueva colección), mostrar los 5 actores y los géneros en los que han participado con más número de géneros diferentes, se debe mostrar el número de géneros diferentes que ha interpretado. Importante! Se necesita previamente filtrar para descartar aquellos actores llamados "Undefined". Aclarar que no se eliminan de la colección, sólo que filtramos para que no aparezcan.

```
1
   var query1 = {cast: {$ne: "Undefined"} }
2
   var fase1 = {$match: query1}
   var query2 = { _id: "$cast",
                                  generos: {$addToSet: "$genres"}}
   var fase2 = {$group: query2}
5
   var query3 = {_id: 1, numgeneros: {$size: "$generos"}, generos: 1}
   var fase3 = {$project: query3}
   var query4 = {numgeneros: -1}
   var fase4 = {$sort: query4}
9
   var fase5 = {$limit: 5}
10
   var etapas = [fase1, fase2, fase3, fase4, fase5]
   db.genres.aggregate(etapas)
```

```
/* 3 */
    "_id" : "Dennis Quaid",
                                                                      "_id" : "James Mason",
    "generos" : [
                                     "_id" : "Michael Caine",
                                                                       "generos" : [
         "Drama",
                                      generos" : [
                                                                          "Romance"
         "Sports",
                                         "Horror".
                                                                           "Science Fiction",
        "Fantasy"
                                         "Romance"
                                                                           "Drama",
        "Science Fiction",
                                         "Science Fiction",
                                                                           "Fantasy"
        "Animated",
                                         "Drama",
                                                                           "Animated",
         "Satire",
                                         "Animated".
                                                                           "Biography",
         "Dance",
                                         "Biography",
                                                                           "Comedy",
        "Biography",
                                         "Comedy",
                                                                           "Action",
                                         "Disaster"
         "Comedy",
                                                                           "Mystery"
         "Disaster'
                                         "Action",
                                                                           "Crime".
                                         "Superhero",
         "Crime",
                                                                           "Musical"
                                         "Mystery",
        "Family",
                                                                           "Undefined",
                                         "Crime",
        "Action",
                                                                           "Western",
                                         "Family",
         "Western",
                                                                           "Noir",
                                          "Undefined".
         "Musical".
                                                                           "Short".
                                          "Thriller",
        "Adventure"
                                                                           "Thriller"
                                          "Suspense",
         "Suspense",
                                                                           "Adventure",
                                         "Adventure"
        "Thriller",
                                                                           "Suspense",
                                         "Spy",
         "Horror",
                                         "War'
                                                                           "War'
         "Romance
    ],
                                      'numgeneros" : 19
                                                                      "numgeneros" : 19
    "numgeneros" : 20
١.
                     "_id" : "Helen Mirren",
                                                    "_id" : "Gene Hackman",
                     "generos" : [
                                                     generos" : [
                         "Spy",
                                                        "Spy",
"Drama",
                        "Fantasy",
                         "Drama",
                                                         "Sports",
                         "Science Fiction",
                                                        "Science Fiction",
                         "Animated",
                                                        "Animated",
                         "Biography",
                                                        "Biography",
                         "Comedy",
                                                        "Disaster"
                         "Erotic"
                                                        "Comedy",
                         "Mystery",
                                                        "Crime"
                         "Crime",
                                                        "Superhero",
                         "Action"
                                                        "Action"
                         "Political",
                                                        "Mystery"
                         "Family".
                                                         "Western".
                         "Adventure"
                                                        "Noir".
                         "Historical",
                                                         "Adventure".
                         "Romance",
                                                         "Suspense",
                         "Horror",
                                                        "Thriller",
                         "Thriller
                                                        "War"
                     "numgeneros" : 18
                                                     'numgeneros" : 18
```

Figura 22: Resultado 22

},

23. Sobre genres (nueva colección), mostrar las 5 películas y su año correspondiente en los que más géneros diferentes han sido catalogados, mostrando esos géneros y el número de géneros que contiene.

```
// 23.
   var query1 = {cast: {$ne: "Undefined"} }
2
   var fase1 = {$match: query1}
3
   var query2 = { _id: {title:"$title", year: "$year"}, generos: {
4
       $addToSet: "$genres"}}
   var fase2 = {$group: query2}
5
   var query3 = {_id: 1, numgeneros: {$size: "$generos"}, generos:1}
6
   var fase3 = {$project: query3}
   var query4 = {numgeneros: -1}
   var fase4 = {$sort: query4}
   var fase5 = {$limit: 5}
10
   var etapas = [fase1, fase2, fase3, fase4, fase5]
11
   db.genres.aggregate(etapas)
```

```
"_id" : {
    "title" : "The Dark Tower",
             "year" : 2017
                                                         "year" : 2017
         generos" : [
              "Action"
                                                        'aeneros" : [
                                                                                                         "_id" : {
| "title" : "My Little Pony: The Movie",
| "year" : 2017
             "Historical",
                                                           "Western",
"Science Fiction",
              "Thriller",
             "Comedy",
"Biography",
"Crime",
"Drama"
                                                            "Action",
                                                            "Adventure".
                                                            "Horror'
                                                           mgeneros" : 6
5
7 },
                                                                           /* 5 */
                          /* 4 */
                                                                                 "_id" : {
                               "_id" : {
    "title" : "Wonder Woman",
                                                                                       "title" : "Dunkirk",
                                                                                      "year" : 2017
                                    "year" : 2017
                                generos" : [
                                     "Superhero",
                                                                                      "Thriller"
                                    "Fantasy",
                                                                                       "Drama",
                                    "War",
"Adventure",
                                                                                      "War",
"Historical",
                                    "Drama",
"Action"
                                                                                      "Adventure
                                                                                  "numgeneros" : 6
```

Figura 23: Resultado 23

24. Sobre genres, mostrar el actor que mas ha participado en las películas de cada uno de los géneros. El actor mas representativo de cada uno de ellos. No se deben contar los genres y actores "Undefined"

```
var fase2 = {$group: query2};
var fase3 = {$sort: {apariciones: -1}}
var query3 = {_id: "$_id.genre", actoresunicos: { $addToSet: "$_id.actor" }, actor_mas_participativo: { $first: "$_id.actor" },
maxapariciones: { $max: "$apariciones" }}
var fase4 = {$group: query3};
var query4 = {_id: 1, actor_mas_participativo: 1, maxapariciones: 1}
var fase5 = {$project: query4};
var query5 = {maxapariciones: -1};
var fase6 = {$sort: query5};
var etapas = [fase1, fase2, fase3, fase4, fase5, fase6];
db.genres.aggregate(etapas);
```

_id * \$	actor_mas_participativo	maxapariciones \$
1 Comedy	Harold Lloyd	186
2 Western	Hoot Gibson	131
3 Animated	Tom and Jerry	85
4 Short	The Three Stooges	64
5 Drama	Bette Davis	56
6 Musical	Bing Crosby	41
7 Adventure	Johnny Weissmuller	29
8 Horror	Boris Karloff	28
9 Action	Jean-Claude Van Damme	22
0 Crime	Edward G. Robinson	21
11 Mystery	Warner Oland	16
2 Romance	Gary Cooper	15
3 Thriller	Nicolas Cage	14
4 Noir	Edward G. Robinson	14
15 War	John Wayne	12
16 Fantasy	Robbie Coltrane	9
17 Science Fiction	William Shatner	9
18 Superhero	Chris Evans	8
9 Biography	Ed Harris	8
20 Family	Maggie Smith	7

	_id * \$	actor_mas_participativo \$	maxapariciones
21	Documentary	Iraq War	5
22	Spy	Dean Martin	5
23	Historical	Tyrone Power	5
4	Suspense	Cary Grant	4
25	Disaster	George Kennedy	4
6	Erotic	Jamie Gillis	4
7	Sports	Elyse Knox	4
28	Satire	Parker Posey	3
9	Political	Jeffrey Wright	3
80	Dance	Adam G. Sevani	3
1	Martial Arts	Jackie Chan	3
32	Performance	Jonas Brothers	2
33	Teen	Sky Hirschkron	2
14	Slasher	Robert Englund	2
35	Live Action	Fred Armisen	2
6	Supernatural	Katie Featherston	2
37	Legal	Marshall Bell	1
88	Silent	Ralph Ince	1
9	Independent	Sarah Prikryl	1
0	Sport	John Wayne	1

Figura 24: Resultado 24

25. Determina el top 3 de géneros con más películas en los últimos 10 años, registrados en la base de datos, y comparalos con el top 3 de géneros de los 10 primeros años registrados

```
//25.
   var Ultimoano = 2018
   var Primerano = 1900
   var Primeradecada = { $gte: Primerano, $lte: Primerano + 9 }
   var Ultimadecada = { $gte: Ultimoano - 9, $lte: Ultimoano }
   var query1 = { genres: { $ne: "Undefined" } }
   var fase1 = { $match: query1 }
10
   // Pipeline para los primeros 10 anios
  var query2a = { year: Primeradecada }
11
  var fase2a = { $match: query2a }
   var query3a = { _id: "$genres", Numpelis: { $sum: 1 } }
13
   var fase3a = { $group: query3a }
14
   var query4a = { Numpelis: -1 }
15
   var fase4a = { $sort: query4a }
16
   var fase5a = { $limit: 3 }
18
   //Pipeline para los ultimos 10 anios
19
  | var query2b = { year: Ultimadecada }
20
  var fase2b = { $match: query2b }
21
   var query3b = { _id: "$genres", Numpelis: { $sum: 1 } }
   var fase3b = { $group: query3b }
23
   var query4b = { Numpelis: -1 }
   var fase4b = { $sort: query4b }
   var fase5b = { $limit: 3 }
   // Ejecutar ambos pipelines usando $facet
28
29
   var etapas = [
       fase1,{
30
31
           $facet: {
               Primeradecada: [fase2a, fase3a, fase4a, fase5a],
32
33
               Ultimadecada: [fase2b, fase3b, fase4b, fase5b] } }]
34
  db.genres.aggregate(etapas)
```



Figura 25: Resultado 25

26. Determinar los 5 géneros a los que pertenecen más películas que tengan en su título la palabra "war".

```
//26.
var query1 = { _id : { title : "$title", genre :"$genres"}}
var fase1 = { $group : query1}
// Filtramos por los titulos que contengan la palabra war
var query2 = {"_id.title" :{ $regex :"war",$options :"i"}}
var fase2 = { $match : query2}
var query3 = { _id: "$_id.genre", total:{ $sum : 1}}
var fase3 = { $group : query3}
var etapas = [fase1, fase2, fase3]
db.genres.aggregate(etapas).sort({total: -1}).limit(5)
```

Figura 26: Resultado 26