UCM

Máster Big Data y Business Analytics

Proyecto: Bases de Datos No SQL

Catedratico: Álvaro Bravo

Presentado por

Skarleth Melissa Motiño Flores

skarlethgm@gmail.com

smotino@ucm.edu



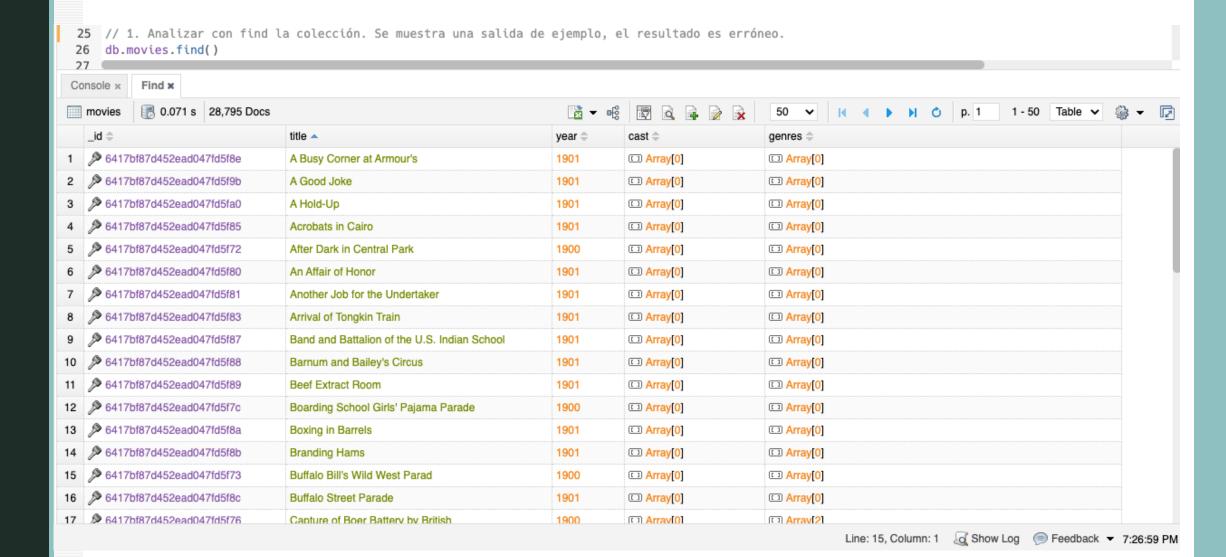
0 – Import dataset Movies:

```
1 - const contents = [
2 -
           content: "/Users/admin/Library/Application Support/NoSQLBooster for MongoDB/databases/movies.json",
           collection: "movies",
           idPolicy: "overwrite_with_same_id",
           //overwrite_with_same_id|always_insert_with_new_id|insert_with_new_id_if_id_exists|skip_documents_with_existing_id|abort_if_id_already_exists|drop_collection_first|log_errors
           //Use the transformer to customize the import result
8 -
           //transformer: (doc)=>{ //async (doc)=>{
                doc["importDate"]= new Date();
9
                doc["oid"] = mb.convert({input:doc["oid"], to:"objectId", onError:"remain_unchanged", onNull:null}); //to: double|string|objectId|bool|date|int|long|decimal
10
11
                return doc; //return null skips this doc
12
           //}
13
14
15
16 - mb.importContent({
17
        connection: "localhost",
18
       database: "clases",
19
       fromType: "file",
                                                                                                                                        Open Connections
20
       batchSize: 2000.
21
        contents
                                                                                                                                        22 })
                                                                                                                                          23
                                                                                                                                          △ III clases (4)
                                                                                                                                            inventory (9)
                                                                                                                                            movies (28.8K)
                                                                                                                                              users
```

Dataset Movies agregado:

local users

1 - Analizar con find la colección.



2 - Contar cuántos documentos (películas) tiene cargado:



3 - Insertar una película:

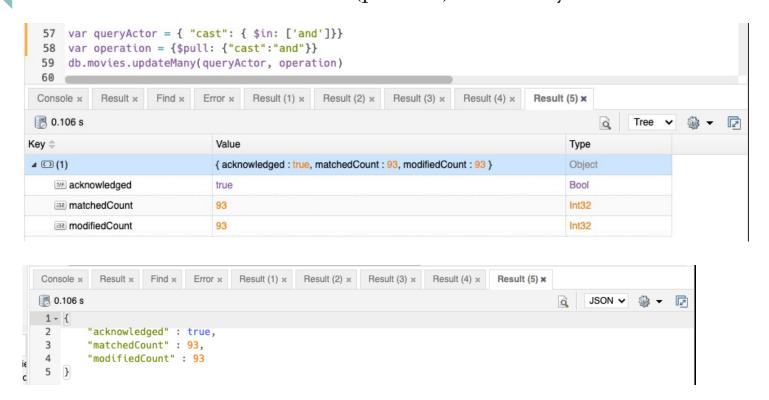
4 - Borrar la película insertada en el punto anterior (en el 3):



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



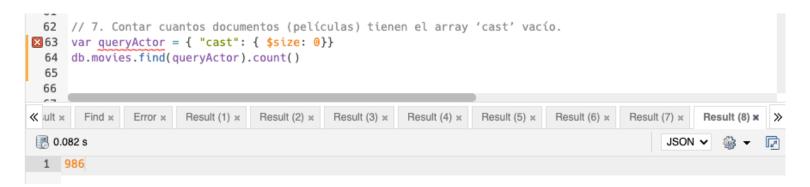
6. Actualizar los documentos cuyo actor (cast) tenga por error el valor "and" como si realmente fuera un actor. Para ello, se debe sacar únicamente ese valor del array cast. Por lo tanto, no se debe eliminar ni el documento (película) ni su array cast con el resto de actores.



Comprobacion, repitiendo la consulta del numeral 5, el resultado es 0:



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



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. Cont. Comprobacion de que se mantiene la estructura array en Cast.

```
75 var queryCastU = { "cast": "Undefined" }
76 db.movies.find(queryCastU)
77
          Result (3) x Result (4) x Result (5) x Result (6) x Result (7) x
        1 /* 1 createdAt:3/19/2023, 7:05:59 PM*/
2 - {
       "_id" : ObjectId("6417bf87d452ead047fd5f71"),
       "title" : "Caught",
       "year" : 1900,
       "cast" : [ "Undefined" ],
       "genres" : [ ]
10 /* 2 createdAt:3/19/2023, 7:05:59 PM*/
11 - {
       "_id" : ObjectId("6417bf87d452ead047fd5f72"),
       "title": "After Dark in Central Park",
13
       "year" : 1900,
15
       "cast" : [ "Undefined" ],
16
       "genres" : [ ]
17 },
```

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



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"].



Comprobar la estructura es aun tipo array

11. Mostrar el año más reciente / actual que tenemos sobre todas las películas. Se muestra una salida de ejemplo, el resultado es erróneo.

```
101 db.movies.find().sort({year:-1}).limit(1)
 102

    Result (7) × Result (8) × Result (9) × Result (10) × Find (1) × Result (11) × Find (2) × Find (3) × Error (2) × Find (4) ×

  1 - {
          "_id" : ObjectId("6417bf89d452ead047fdceff"),
          "title" : "Insidious: The Last Key",
          "vear" : 2018.
          "cast" : [
              "Lin Shave".
              "Angus Sampson"
              "Leigh Whannell"
              "Spencer Locke".
 10
              "Caitlin Gerard"
 11
              "Kirk Acevedo".
 12
              "Bruce Davison"
 13
 14
          "genres" : [ "Horror", "Thriller" ]
 15 }
```

• 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.



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.

```
119  var fase1 = { $group: { "_id": "$year", "total": { $sum: 1 } } }

120  var fase2 = { $match: { "_id": { $gte: 1960 } } }

121  var fase3 = { $match: { "_id": { $lte: 1969 } } }

122  var fase4 = { $group: { "_id": null, "total": { $sum: "$total" } } }

123  db.movies.aggregate([fase1, fase2, fase3, fase4])

124

125

125

126  var fase4 = { $group: { "_id": null, "total": { $sum: "$total" } } }

127

128

129

120  var fase4 = { $group: { "_id": null, "total": { $sum: "$total" } } }

129

120  var fase4 = { $group: { "_id": null, "total": { $sum: "$total" } } }

120  var fase2 = { $match: { "_id": fase3 = { $match: { "_id": null, "total": fase4 = { $sum: "$total" } } }

120  var fase2 = { $match: { "_id": fase3 = { $match: { "_id": fase4 = { $sum: "$total": } } } }

120  var fase4 = { $group: { "_id": null, "$total": fase4 = { $sum: fa
```

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. →

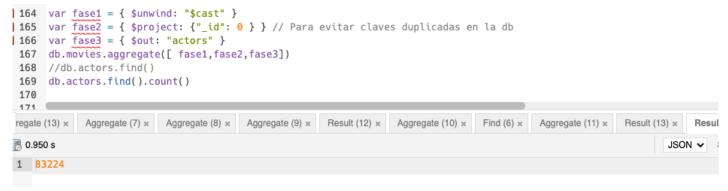
```
129 var YearArray = { "year": "$year" }
 130 //Agrupar por año
131 var fase1 = { $group: { "_id": "$year", "numero_peliculas": { $sum: 1 } } }
132 //renombrar id por año
1133 var fase2 = { $project: { "_id": 0, "year": "$_id", "numero_peliculas": 1 } }
     var fase3 = { $group: { _id: "$numero_peliculas", "number_years": { $sum: 1 }, "items": {$push: YearArray } } }
135 //Agrupar por numero de películas, para encontrar todos los años que tienen en mismo numero de películas.
136 //renombrar _id por numero de peliculas
137 var fase4 = { $project: { "_id": 0, "numero_peliculas": "$ id", "years": "$items" } }
138 //Ordenar descendentemente y limitar a 1 para mostrar solo el maximo numero de peliculas,
139 //con los años que correspondan a ese mismo # de peliculas.
1140 var fase5 = { $sort: { "numero_peliculas": -1 } }
141 var fase6 = { $limit: 1 }
142 var fase7 = { $project:{"_id_year":"$years.year" , "# pelis":"$numero_peliculas"} }
143 db.movies.aggregate([fase1, fase2, fase3, fase4, fase5, fase6, fase7])
 144
Aggregate (1) × Aggregate (2) × Aggregate (3) × Aggregate (4) × Aggregate (5) ×
movies 0.088 s 1 Doc
                                                                            ◆ ▶ Ø p. 1 1-1 JSON ✔ 🚳 ▼
        " id year" : [ 1919 ],
        "# pelis" : 634
```

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.

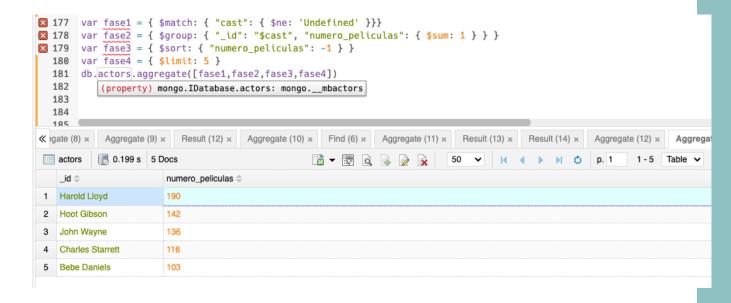
En este caso, aparecen 3 años listados con el menor numero de peliculas por año.



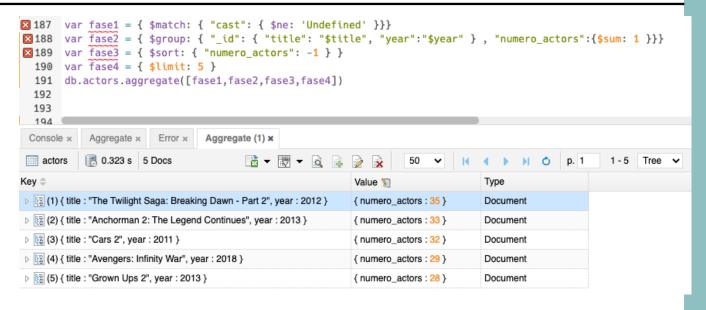
16. Guardar en nueva colección llamada "actors" realizando la fase \$unwind por actor. Después, contar cuantos documentos existen en la nueva colección.



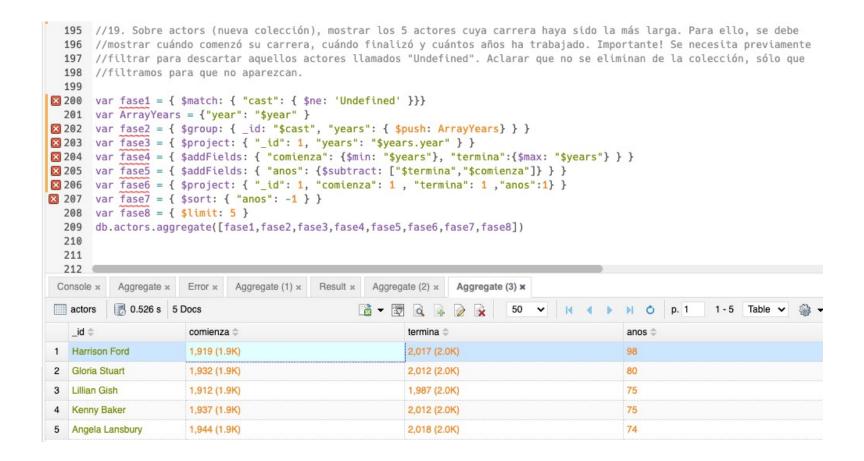
17. Sobre actors (nueva colección), mostrar la lista con los 5 actores que han participado en más películas mostrando el número de películas en las que ha participado. Importante!
Se necesita previamente filtrar para descartar aquellos actores llamados "Undefined". Aclarar que no se eliminan de la colección,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



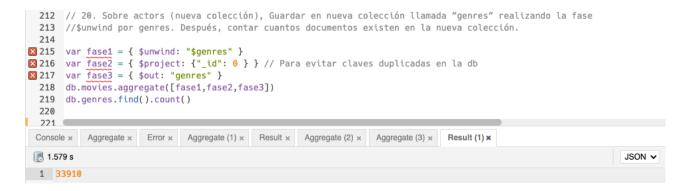
18. Sobre actors (nueva colección), agrupar por película y año mostrando las 5 en las que más actores hayan participado, mostrando el número total de actores.



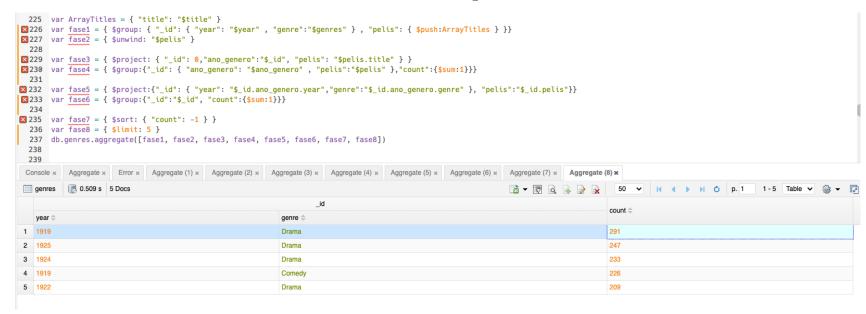
19. Sobre actors (nueva colección), mostrar los 5 actores cuya carrera haya sido la más larga. Para ello, se debe mostrar cuándo comenzó su carrera, cuándo finalizó y cuántos años ha trabajado. Importante! Se necesita previamente filtrar para descartar aquellos actores llamados "Undefined". Aclarar que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



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.



• 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.



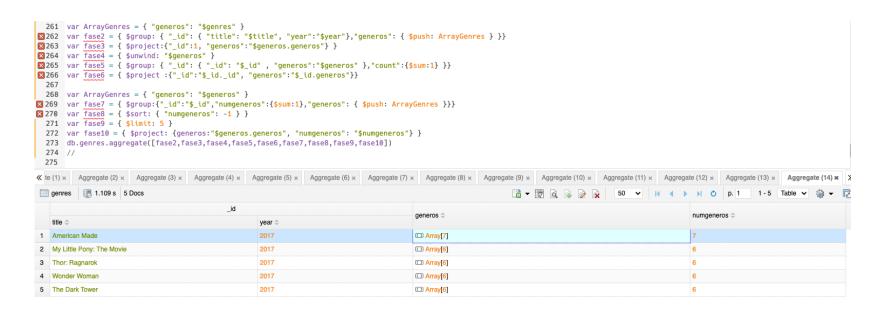
22. Sobre genres (nueva colección), mostrar los 5 actores y los géneros en los que han participado con más número de géneros diferentes, se debe mostrar el número de géneros diferentes que ha interpretado. Importante! Se necesita previamente filtrar para descartar aquellos actores llamados "Undefined".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



Salida en formato Json:

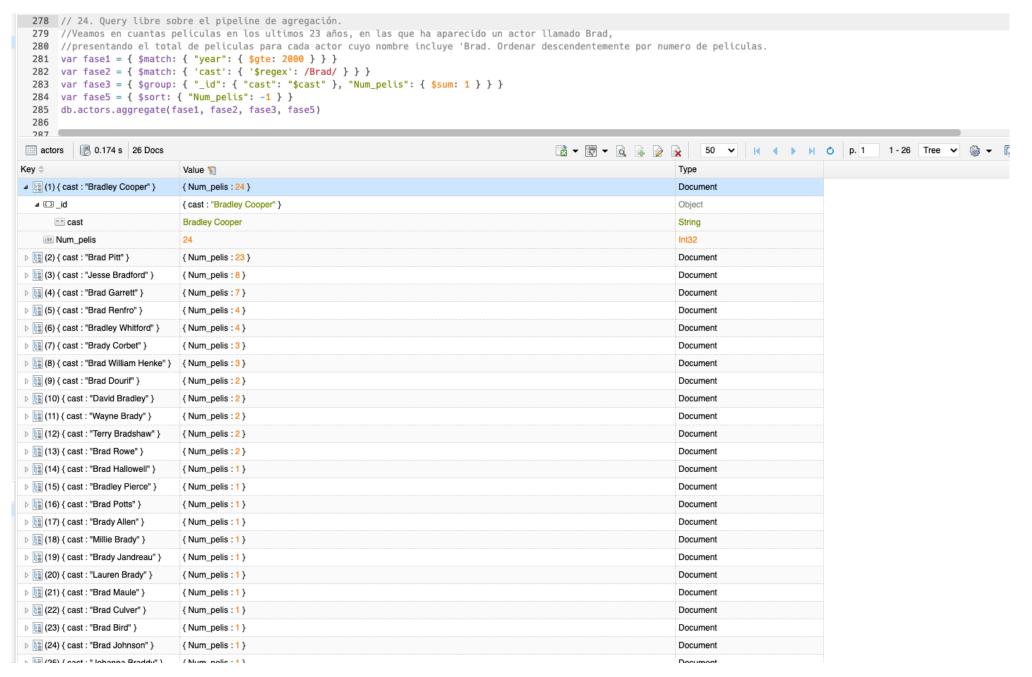
```
83 /* 4 */
                                                                                                                                                                            109 /* 5 */
                                                                                     56 /* 3 */
                                                29 /* 2 */
                                                                                                                                 84 - {
genres 0.726 s 5 Docs
                                                                                                                                                                            110 - {
                                                                                     57 - {
                                                30 - {
                                                                                                                                          " id" : "Gene Hackman",
                                                                                                                                 85
 1 /* 1 */
                                                                                                                                                                            111
                                                                                                                                                                                      "_id" : "Helen Mirren",
                                                                                     58
                                                                                             "_id" : "Michael Caine",
                                                31
                                                       " id" : "James Mason",
                                                                                                                                          "numgeneros": 18,
 2 - {
                                                                                                                                                                            112
                                                                                                                                                                                      "numgeneros" : 18,
                                                                                     59
                                                                                             "numgeneros": 19.
                                                32
                                                        "numgeneros": 19,
        " id" : "Dennis Quaid",
                                                                                                                                 87 -
                                                                                                                                          "generos" : [
                                                                                             "generos" : [
                                                                                                                                                                            113 -
                                                                                                                                                                                      "generos" : [
                                                33 -
                                                        "generos" : [
         "numgeneros": 20,
                                                                                                                                 88
                                                                                                                                              "Noir",
                                                                                     61
                                                                                                 "Horror",
                                                34
                                                           "Short",
                                                                                                                                                                            114
                                                                                                                                                                                           "Historical",
         "generos" : [
                                                                                     62
                                                                                                                                 89
                                                                                                                                              "War",
                                                                                                 "Adventure",
                                                35
                                                           "War",
            "Disaster",
                                                                                                                                                                            115
                                                                                                                                                                                           "Adventure",
                                                36
                                                                                     63
                                                                                                                                 90
                                                                                                                                               "Spy",
                                                           "Romance".
                                                                                                 "Family",
            "Biography",
                                                                                                                                                                            116
                                                                                                                                                                                           "Drama".
                                                37
                                                                                     64
                                                           "Drama",
                                                                                                                                 91
                                                                                                 "Biography"
                                                                                                                                              "Adventure".
            "Thriller",
                                                                                                                                                                            117
                                                                                                                                                                                           "Romance",
                                                38
                                                           "Mystery",
                                                                                     65
                                                                                                 "Disaster",
                                                                                                                                 92
            "Action",
                                                                                                                                              "Action",
                                                                                                                                                                            118
                                                                                                                                                                                           "Science Fiction".
                                                39
                                                           "Thriller",
                                                                                     66
                                                                                                 "Spy",
10
            "Sports",
                                                                                                                                 93
                                                                                                                                               "Sports",
                                                                                                                                                                            119
                                                                                                                                                                                           "Crime".
                                                40
                                                           "Action",
11
            "Musical",
                                                                                     67
                                                                                                 "Romance"
                                                                                                                                 94
                                                                                                                                              "Animated".
                                                41
                                                           "Comedy",
                                                                                                                                                                            120
                                                                                                                                                                                           "Fantasy",
12
            "Dance",
                                                                                     68
                                                                                                 "Animated"
                                                                                                                                 95
                                                                                                                                              "Biography",
                                                42
13
            "Adventure",
                                                           "Western"
                                                                                                                                                                            121
                                                                                                 "Suspense",
                                                                                                                                                                                           "Mystery",
                                                                                                                                 96
                                                                                                                                              "Disaster",
14
                                                43
            "Family",
                                                           "Animated"
                                                                                     70
                                                                                                 "Mystery",
                                                                                                                                                                            122
                                                                                                                                                                                           "Animated"
                                                                                                                                 97
                                                                                                                                              "Science Fiction",
15
            "Drama",
                                                44
                                                           "Crime".
                                                                                     71
                                                                                                 "War",
                                                                                                                                                                            123
                                                                                                                                                                                           "Action",
16
            "Horror",
                                                45
                                                                                                                                 98
                                                           "Fantasv".
                                                                                                                                              "Thriller",
                                                                                     72
                                                                                                 "Science Fiction",
                                                                                                                                                                            124
                                                                                                                                                                                           "Horror",
17
            "Suspense",
                                                46
                                                           "Suspense",
                                                                                                                                 99
                                                                                                                                              "Drama",
                                                                                     73
                                                                                                 "Comedy",
18
                                                                                                                                                                            125
                                                                                                                                                                                          "Family",
            "Romance",
                                                47
                                                           "Undefined",
                                                                                                                                100
                                                                                                                                              "Suspense",
                                                                                     74
                                                                                                 "Drama",
19
            "Animated"
                                                48
                                                           "Science Fiction",
                                                                                                                                                                            126
                                                                                                                                                                                           "Biography"
                                                                                     75
                                                                                                 "Thriller",
                                                                                                                                101
                                                                                                                                               "Western",
20
            "Western",
                                                49
                                                           "Biography",
                                                                                                                                                                            127
                                                                                                                                                                                           "Political"
                                                                                     76
                                                                                                                                102
                                                                                                                                              "Crime",
                                                                                                 "Crime",
21
            "Comedy",
                                                50
                                                           "Noir",
                                                                                                                                                                            128
                                                                                                                                                                                           "Comedy",
                                                                                     77
                                                                                                                                103
22
                                                                                                 "Superhero",
                                                                                                                                              "Superhero",
            "Fantasy",
                                                51
                                                           "Musical",
                                                                                                                                                                            129
                                                                                                                                                                                           "Erotic",
                                                                                     78
23
                                                                                                 "Action",
            "Satire",
                                                                                                                                104
                                                52
                                                                                                                                               "Mystery",
                                                           "Adventure"
24
            "Crime",
                                                                                     79
                                                                                                                                                                            130
                                                                                                                                                                                           "Spy",
                                                                                                 "Undefined"
                                                                                                                                105
                                                                                                                                               "Comedy"
                                                53
25
            "Science Fiction"
                                                                                                                                                                                           "Thriller"
                                                                                     80
                                                                                                                                                                            131
                                               54 },
                                                                                                                                106
26
                                                                                     81 },
                                                                                                                                                                            132
                                                                                                                                107 },
27 },
                                                                                                                                                                            133 }
                                                                                                                                108
28
```

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.

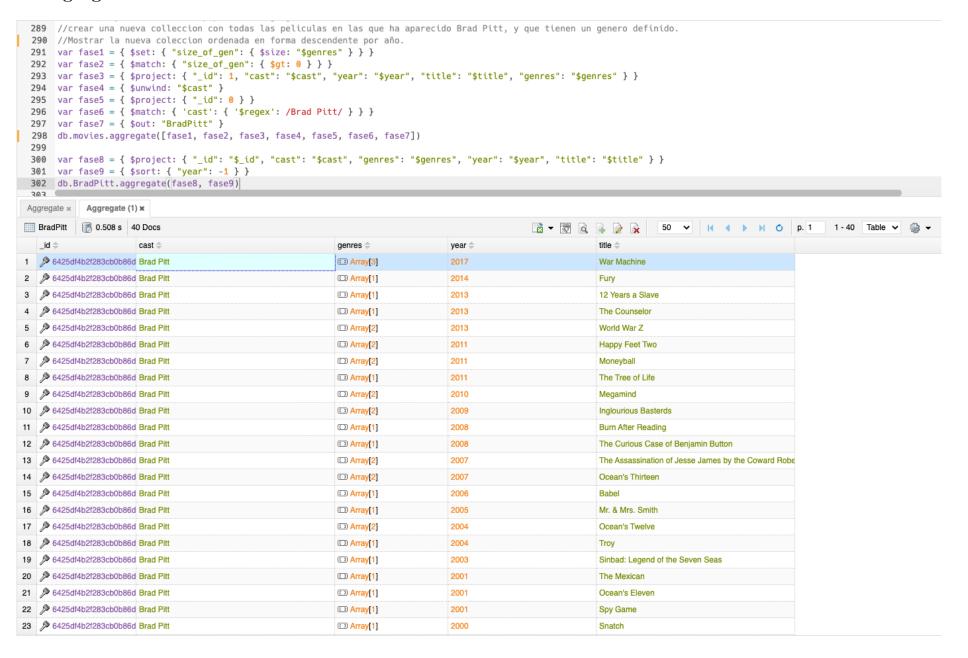


```
1 /* 1 */
                                                                                               36 /* 3 */
                                        19 /* 2 */
2 - {
                                                                                                                                           53 /* 4 */
                                                                                               37 - {
                                                                                                                                                                                    70 /* 5 */
                                        20 - {
                                                                                                                                           54 - {
                                                                                                        " id" : {
                                                                                                                                                                                    71 - {
                                        21 -
                                                " id" : {
            "title": "American Made"
                                                                                                                                           55 -
                                                                                                                                                   " id" : {
                                                                                                                                                                                    72 -
                                                                                                                                                                                             " id" : {
                                        22
                                                                                                            "title": "Thor: Ragnarok",
                                                    "title" : "My Little Pony: The Movie",
            "vear" : 2017
                                                                                                                                           56
                                                                                                                                                        "title": "Wonder Woman",
                                                                                                                                                                                    73
                                                                                                                                                                                                 "title": "The Dark Tower",
                                        23
                                                    "vear" : 2017
                                                                                               40
                                                                                                            "year" : 2017
                                                                                                                                           57
                                                                                                                                                        "year" : 2017
                                                                                                                                                                                    74
                                                                                                                                                                                                 "year" : 2017
                                                                                               41
                                                                                                       },
                                        24
                                                },
        "generos" : [
                                                                                                                                           58
                                                                                                                                                   },
                                                                                                                                                                                    75
                                                                                                                                                                                             },
                                                                                               42 -
                                                                                                        "generos" : [
                                        25 -
                                                "generos" : [
            "Drama".
                                                                                                                                           59 -
                                                                                                                                                    "generos" : [
                                                                                               43
                                                                                                            "Superhero",
                                                                                                                                                                                    76 -
                                                                                                                                                                                             "generos" : [
                                        26
                                                    "Comedy",
            "Comedy",
                                                                                                                                                        "Action"
                                                                                                                                                                                    77
                                                                                                                                                                                                 "Action",
                                                                                               44
                                                                                                            "Science Fiction",
                                        27
                                                    "Adventure"
            "Action".
                                                                                                                                           61
                                                                                                                                                        "Fantasy",
                                                                                                                                                                                    78
                                                                                                                                                                                                 "Science Fiction",
                                        28
                                                    "Fantasy",
                                                                                               45
                                                                                                            "Comedy",
11
            "Historical",
                                                                                                                                           62
                                                                                                                                                        "Drama",
                                                                                                                                                                                    79
                                                                                                                                                                                                 "Adventure",
                                        29
                                                                                               46
                                                                                                            "Fantasy",
                                                    "Family",
12
            "Crime".
                                                                                                                                           63
                                                                                                                                                        "Superhero",
                                                                                                                                                                                                 "Horror",
                                                                                               47
                                                                                                            "Adventure".
                                        30
                                                    "Animated",
13
            "Biography",
                                                                                                                                                        "Adventure",
                                                                                               48
                                                                                                            "Action"
                                                                                                                                                                                    81
                                                                                                                                                                                                 "Fantasy",
                                        31
                                                    "Musical"
14
            "Thriller"
                                                                                                                                           65
                                                                                                                                                        "War"
                                                                                                                                                                                    82
                                                                                                                                                                                                 "Western"
                                                                                               49
                                        32
15
                                                                                                                                           66
                                                                                                                                                                                    83
                                        33
                                                                                               50
                                                                                                        "numgeneros" : 6
                                                 "numgeneros" : 6
        "numgeneros" : 7
                                                                                                                                           67
                                                                                                                                                    "numgeneros" : 6
                                                                                               51 },
                                                                                                                                                                                             "numgeneros" : 6
                                        34 },
17 },
                                                                                                                                           68
                                                                                                                                                                                    85 }
                                                                                               52
```

24. Ejercicios libre de agregacion



25. Ejercicios libre de agregacion



26. Ejercicios libre de agregacion

