



<> Code

Issues

Pull requests

Actions

Projects

Wiki

Security



main ▾

ud2-practica-3-mongo-db-Dansarasix-DML / Readme.md



Dansarasix-DML update changes

7d96073 · 8 months ago



324 lines (261 loc) · 8.17 KB

Preview

Code

Blame



Raw



Big Data Aplicado

UD2 - Procesado y Presentación de Datos Almacenados

🔗 Práctica 3 MongoDB

Recuerda que para hacer la prácticas puedes optar por cualquiera de la las 3 opciones.

- Instalarla en tu máquina local
- Usar [Atlas MongoDB](#)
- Crear un contenedor docker.

Sobre la Base de Datos **sample-mflix** y usando la Collection **movies** realiza las siguientes operaciones:

1. Encuentra todas las películas con información completa que tengan una duración superior a 120 minutos.

Usamos `find()` para buscar aquellas películas que tengan un *runtime* superior a 120.

```
> db["movies"].find( {runtime: { $gt:120 }} ).count()  
< 3313
```

```
db["movies"].find( {runtime: { $gt:120 }} )
```



2. Encuentra todas las películas con información completa que tengan el género "Short".

Usando de nuevo `find()` en este caso sobre el array `genres` para buscar un elemento que sea "Short".

```
> db["movies"].find( { genres: "Short" } ).count()  
< 442
```

```
db["movies"].find( { genres: "Short" } )
```



3. Lista todas las películas con información completa y han recibido más de 1000 votos en IMDb.

Pues hacemos igual que antes pero sobre el array `imdb` en el campo `votes`.

```
> db["movies"].find( {"imdb.votes": { $gt:1000 }} ).count()  
< 13289
```

```
db["movies"].find( {"imdb.votes": { $gt:1000 }} )
```



4. Lista todas las películas y tengan una calificación de IMDb superior a 7.

Lo mismo que antes pero sobre el campo `rating`.

```
> db["movies"].find( {"imdb.rating": { $gt:7 }} ).count()  
< 8160
```

```
db["movies"].find( {"imdb.rating": { $gt:7 }} )
```



5. Recupere todas las películas y una calificación de espectador superior a 4 en Tomatoes.

```
> db["movies"].find( {"tomatoes.critic.rating": { $gt:4 }} ).count()  
< 9548
```

```
db["movies"].find( {"tomatoes.critic.rating": { $gt:4 }} )
```



6. Lista todas las películas que hayan recibido al menos un premio.

```
> db["movies"].find( {"awards.wins": { $gt:1 }} ).count()  
< 10751
```

```
db["movies"].find( {"awards.wins": { $gt:1 }} )
```



7. Muestra las películas sólo con título, idiomas, estreno, directores, escritores, premios, año, géneros, duración, reparto, países que tengan al menos una nominación.

```
> db["movies"].find( {"awards.nominations": { $gt:1 }} ).count()  
< 12052
```

```
db["movies"].find( {"awards.nominations": { $gt:1 }} )
```



8. Muestra las películas sólo con título, idiomas, estreno, directores, escritores y países de la colección de 'películas' que se estrenaron después del 2014.

```
> db["movies"].find( {$expr: {  
    $gt: [{ $year: "$released" }, 2014]  
}} ).count()  
< 783
```

```
db["movies"].find( { $expr: {  
    $gt: [{ $year: "$released" }, 2014]  
} } )
```



9. Lista las películas sólo con título, idiomas, estreno, directores, escritores y países que tengan la palabra "scene" en el título.

```
> db["movies"].find( { title: { $regex: "scene", $options: "i" } } ).count()  
< 12
```

```
db["movies"].find( { title: { $regex: "scene", $options: "i" } } )
```



10. Lista las películas sólo con título, idiomas, estreno, directores, espectadores, escritores y países que tengan una calificación de espectadores de al menos 3 y menos de 4 en Tomatoes.

```
> db["movies"].find( { $and: [ {"tomatoes.critic.rating": { $gte:3 }}, {"tomatoes.critic.rating": { $lt:4 }} ] } ).count()
< 599
```

```
db["movies"].find( {
  $and: [
    {"tomatoes.critic.rating": { $gte:3 }},
    {"tomatoes.critic.rating": { $lt:4 }}
  ]
})
```



11. Lista las películas sólo con título, idiomas, estreno, año, directores, escritores y países que se estrenaron antes del año 1900.

```
> db["movies"].find( {$expr: {
  $lt: [{ $year: "$released" }, 1900]
}} ).count()
< 471
```

```
db["movies"].find( {$expr: {
  $lt: [{ $year: "$released" }, 1900]
}} )
```



12. Muestra las películas sólo con título, idiomas, estreno, duración, directores, escritores, países que tengan una duración entre 60 y 90 minutos.

```
> db["movies"].find( { $and: [ {runtime: { $gt:60 }}, {runtime: { $lt:90 }} ] } ).count()
< 3858
```

```
db["movies"].find( { $and: [ {runtime: { $gt:60 }}, {runtime: { $lt:90 }} ] } )
```



13. Muestra las películas sólo con título, idiomas, estreno, tiempo de ejecución, directores, escritores, países, imdb para las 5 películas con las calificaciones más altas de IMDb.

ALL RESULTS

OUTPUT OPTIONS ▾

Showing 1 - 5 [count results](#)

VIEW



```
title: "Band of Brothers"
calificacion_IMDB: 9.6
```

```
title: "Planet Earth"
calificacion_IMDB: 9.5
```

```
title: "A Brave Heart: The Lizzie Velasquez Story"
calificacion_IMDB: 9.4
```

```
title: "The Civil War"
calificacion_IMDB: 9.4
```

```
title: "The Civil War"
calificacion_IMDB: 9.4
```

```
db.getCollection('movies').aggregate(
[
  { $match: { 'imdb.rating': { $lte: 10 } } },
  //{ $match: { 'imdb.rating': { $ne: null } } },
  { $sort: { 'imdb.rating': -1 } },
  { $limit: 5 },
  {
    $project: {
      _id: 0,
      title: 1,
      calificacion_IMDB: '$imdb.rating'
    }
  }
],
{ maxTimeMS: 60000, allowDiskUse: true }
);
```



14. Lista las películas con el tiempo de ejecución promedio de las películas estrenadas en cada país.

ALL RESULTS

OUTPUT OPTIONS ▾

Showing 1 - 20 [count results](#)

VIEW



```
country: "India"
averageRuntime: 142
```

```
country: "Belarus"
averageRuntime: 132.5
```

```
country: "Malta"
averageRuntime: 128.8
```

```
country: "Montenegro"
averageRuntime: 125.5
```

```
country: "Algeria"
averageRuntime: 123.31
```

```
db.getCollection('movies').aggregate(  
  [  
    { $unwind: { path: '$countries' } },  
    {  
      $group: {  
        _id: '$countries',  
        averageRuntime: { $avg: '$runtime' }  
      }  
    },  
    {  
      $project: {  
        country: '$_id',  
        _id: 0,  
        averageRuntime: {  
          $round: ['$averageRuntime', 2]  
        }  
      }  
    },  
    { $sort: { averageRuntime: -1 } }  
  ],  
  { maxTimeMS: 60000, allowDiskUse: true }  
);
```



15. Muestra el género más común entre las películas y cuenta cuantas veces aparece.

ALL RESULTS

OUTPUT OPTIONS ▼

```
genre : "Drama"  
count : 12385
```

```
db.getCollection('movies').aggregate(  
  [  
    { $unwind: { path: '$genres' } },  
    {  
      $group: {  
        _id: '$genres',  
        count: { $sum: 1 }  
      }  
    },  
    {  
      $project: {  
        _id: 0,  
        genre: '$_id',  
        count: '$count'  
      }  
    }  
  ]
```



```

    }
  },
  { $sort: { count: -1 } },
  { $limit: 1 }
],
{ maxTimeMS: 60000, allowDiskUse: true }
);

```

16. Encuentra los/as 10 directores/as con más películas

ALL RESULTS **OUTPUT OPTIONS** Showing 1 – 10 [count results](#) < > VIEW  

```

director : "Woody Allen"
count : 40

```

```

director : "Martin Scorsese"
count : 32

```

```

director : "Takashi Miike"
count : 31

```

```

director : "Sidney Lumet"
count : 29

```

```

director : "John Ford"
count : 29

```

```

db.getCollection('movies').aggregate(
[
  { $unwind: { path: '$directors' } },
  {
    $group: {
      _id: '$directors',
      count: { $sum: 1 }
    }
  },
  {
    $project: {
      _id: 0,
      director: '$_id',
      count: '$count'
    }
  },
  { $sort: { count: -1 } },
  { $limit: 10 }
],
{ maxTimeMS: 60000, allowDiskUse: true }
);

```



17. Desarrolla una consulta para encontrar la calificación promedio de IMDb para películas con diferentes calificaciones (por ejemplo, 'PG', 'R', 'G')

ALL RESULTS

OUTPUT OPTIONS ▾

Showing 1 – 18 [count results](#)

VIEW



```
type : "Not Rated"
AVGIMDBRating : 7.7
```

```
type : "TV-MA"
AVGIMDBRating : 7.36
```

```
type : "PASSED"
AVGIMDBRating : 7.31
```

```
type : "OPEN"
AVGIMDBRating : 7.2
```

```
type : "APPROVED"
AVGIMDBRating : 7.17
```

```
db.getCollection('movies').aggregate(
[
  {
    $group: {
      _id: '$rated',
      AVGIMDBRating: { $avg: '$imdb.rating' }
    }
  },
  {
    $project: {
      _id: 0,
      type: '$_id',
      AVGIMDBRating: {
        $round: ['$AVGIMDBRating', 2]
      }
    }
  },
  { $sort: { AVGIMDBRating: -1 } }
],
{ maxTimeMS: 60000, allowDiskUse: true }
);
```



18. Desarrolla una consulta encontrar la película más antigua ganadora de un premio

ALL RESULTS

OUTPUT OPTIONS ▾

```
title : "The Great Train Robbery"
yearReleased : 1903
premios : 1
```



```
db.getCollection('movies').aggregate(  
  [  
    {  
      $match: {  
        'awards.wins': { $eq: 1 },  
        released: { $ne: null }  
      }  
    },  
    {  
      $project: {  
        _id: 0,  
        title: 1,  
        yearReleased: { $year: '$released' },  
        premios: '$awards.wins'  
      }  
    },  
    { $sort: { yearReleased: 1 } },  
    { $limit: 1 }  
  ],  
  { maxTimeMS: 60000, allowDiskUse: true }  
);
```



19. Encuentra la película con la calificación de IMDb más alta y de ellas, la que tenga la calificación de espectadores más altas en Tomatoes

ALL RESULTS **OUTPUT OPTIONS** ▼

```
title : "The Shawshank Redemption"  
calificacion_IMDB : 9.3  
calificacion_Tomatoes : 8.2
```

```
db.getCollection('movies').aggregate(  
  [  
    { $match: { 'imdb.rating': { $lte: 10 } } },  
    //{ $match: { 'imdb.rating': { $ne: null } } },  
    {  
      $match: {  
        'tomatoes.critic.rating': { $lte: 100 }  
        //'tomatoes.critic.rating': { $ne: null }  
      }  
    },  
    {  
      $sort: {
```



```
        'imdb.rating': -1,
        'tomatoes.critic.rating': -1
    }
},
{ $limit: 1 },
{
    $project: {
        _id: 0,
        title: 1,
        calificacion_IMDB: '$imdb.rating',
        calificacion_Tomatoes:
            '$tomatoes.critic.rating'
    }
}
],
{ maxTimeMS: 60000, allowDiskUse: true }
);
```