

Tuần 4-5-6-7: Aggregation Operations

Mục tiêu:

- Hiểu rõ về quy trình trong nội dung aggregation pipelines.
- Thực hiện các stage: match, project, group,...
- Hiểu rõ về các accumulators và expression trong mỗi stage.

Cú pháp: Aggregation

```
db.collectionname.aggregate([  
    {<Stage 1>},  
    {<Stage 2>},  
    .....  
    {<Stage n>},  
])
```

- [\\$match:](#)

Dùng để lọc các document theo một số điều kiện.

```
db.collectionname.aggregate([{$match:{<expression>}}])
```

- [\\$project:](#)

Cho phép định nghĩa các thuộc tính cần lấy từ collection ban đầu.

```
db.collectionname.aggregate([{$project:{'field1':1,'field':0,  
<field>: <expression>...}}])
```

- [\\$group:](#)

Cho phép gom nhóm các document theo các field của collection ban đầu.

```
db.collectionname.aggregate([  
    $group:{_id:'$field',  
    accumulator:<expression>}}  
])
```

- [\\$sort:](#) sắp xếp các document.
- [\\$skip:](#) bỏ qua n documents và chuyển các documents còn lại
- [\\$limit:](#) chuyển n documents đầu tiên, do đó giới hạn chúng.
- [\\$unwind:](#) dùng cho các documents đang sử dụng arrays để trả về document cho từng phần tử.
- [\\$out:](#) ghi các documents kết quả vào một collection mới.

Expressions: tham chiếu đến tên của field trong input documents

VD: { \$group : { _id : “\$id“, total: { \$sum:”\$fare“} } } trong đó \$id và \$fare là expressions.

Accumulators: chúng được sử dụng trong group stage

- **sum:** tính tổng các giá trị số cho các documents trong mỗi group
- **count:** đếm tổng số documents trong group
- **avg:** tính trung bình của tất cả các giá trị đã cho từ tất cả documents
- **min:** trả về giá trị nhỏ nhất trong tất cả các documents
- **max:** trả về giá trị lớn nhất trong tất cả các documents
- **first:** trả về document đầu tiên trong nhóm
- **last:** trả về document cuối cùng trong nhóm

Bài 1. Collection Sinhvien

1. Đếm số sinh viên của mỗi lớp, thông tin gồm mã lớp và số sinh viên.
2. Tính điểm trung bình của từng lớp, thông tin gồm mã lớp và điểm trung bình.
3. Đếm số sinh viên có điểm trung bình từ 5 trở lên của từng lớp, thông tin gồm mã lớp, số sinh viên.
4. Đếm số sinh viên có điểm trung bình từ 3.0 đến dưới 6.5 của từng lớp, thông tin gồm mã lớp, số sinh viên.
5. Đếm số sinh viên không có số điện thoại của từng lớp, thông tin gồm mã lớp, số sinh viên không điện thoại.
6. Đếm số sinh viên thực tế theo từng lớp, thông tin bao gồm thông tin của lớp học và tổng số sinh viên thực. (gồm 2 collection)
7. Tìm lớp học có tổng số sinh viên thực tế cao nhất. Thông tin bao gồm thông tin của lớp học và tổng số sinh viên. [{ _id: 'DHKTPM15B', soluong: 59 }]
8. Đếm số sinh viên có điểm trung bình từ 9.0 trở lên theo từng lớp, sắp xếp theo tên sinh viên. Thông tin bao gồm thông tin của lớp học và tổng số sinh viên.
9. Xuất danh sách sinh viên có điểm trung bình từ 9.0 trở lên theo từng lớp, sắp xếp theo tên sinh viên ra collection riêng biệt. Thông tin bao gồm thông tin của sinh viên.
10. Đếm số lớp theo chuyên ngành. Thông tin gồm mã chuyên ngành, số lớp
11. Tính tổng số sinh viên theo từng chuyên ngành. Thông tin bao gồm thông tin của chuyên ngành và tổng số sinh viên.

Bài 2. Collection customers

1. Xuất danh sách khách hàng với các thông tin address, city, state.
2. Xuất các khách hàng có city ở “Woburn” và state là “MA”.
3. Xuất thông tin khách hàng bao gồm individual, first_name, last_name của officer.
4. Nhóm danh sách khách hàng theo city và cho biết số lượng khách hàng tương ứng với điều kiện là các khách hàng có city ở “Salem”.
5. Xuất danh sách 5 khách hàng đầu tiên được sắp xếp theo thứ tự tăng dần của postal_code và giảm dần của fed_id.

6. Lọc các khách hàng có state là "MA". Sau đó, thực hiện nhóm các khách hàng vừa tìm được theo city và đếm số lượng khách hàng tương ứng với danh sách city vừa nhóm. Tiếp theo thực hiện xuất 03 dữ liệu cuối cùng được sắp xếp kết quả giảm dần theo postal_code. `***{city}`
7. Xuất danh sách khách hàng có tồn tại field business.
8. Xuất danh sách khách hàng có tồn tại field business với các thông tin fed_id, cust_type_cd, state, khuyếnMai. Trong đó field khuyếnMai được tính như sau: nếu field state = "MA" thì sẽ được hưởng khuyến mãi (khuyếnMai: true), ngược lại sẽ không được hưởng khuyến mãi (khuyếnMai: false).
KM:{ \$cond: { if: {state:"MA"}, then: "giảm 10%", else: "không km" } }
9. Nếu field khuyếnMai là true (câu 8) thì sẽ hiển thị là "Được giảm 10%", ngược lại sẽ hiển thị "Không được hưởng khuyến mãi".

Bài 3. Collection zips

1. Hiển thị n documents từ document thứ k. (n, k tùy ý)
2. Chèn thêm 1 document mới. (tùy ý)
3. Cập nhật thông tin của một document khi biết id bất kỳ.
4. Tìm dân số của thành phố (city) PALMER.
5. Tìm các document có dân số (pop) >100000
6. Tìm dân số của thành phố (city) FISHERS ISLAND
7. Tìm các thành phố có dân số từ 10 – 50
8. Tìm tất cả các thành phố của bang MA có dân số trên 500
9. Tìm tất cả các bang (không trùng)
10. Tìm tất cả các bang mà có chứa ít nhất 1 thành phố có dân số trên 100000
11. Tính tổng số dân (pop) theo từng bang (state).
12. Tìm tất cả các bang có tổng dân số trên 10.000.000
13. Tính dân số trung bình (các thành phố) theo từng bang (state).
14. Tìm những document của bang 'CT' và thành phố 'WATERBURY'
15. Bang WA có bao nhiêu city (nếu trùng chỉ tính 1 lần)**
16. Tính số city của mỗi bang (nếu trùng chỉ tính 1 lần), kết quả giảm dần theo số city
17. Group theo state và city, sum:1; Group state, sum
18. Tìm ra các thành phố có dân số (pop) lớn (nhỏ) nhất.
19. Sort theo pop, \$group☺)
20. Tìm bang có dân số (pop) lớn (nhỏ) nhất.
21. Xuất những document có dân số dưới dân số trung bình của mỗi city

Bài 4. Collection movies

movies

```
{
  _id: ObjectId("573a1390f29313caabcd418c"),
  title: 'The House of the Devil',
  year: 1896,
  runtime: 3,
```

```

cast: [ "Jeanne d'Alcy", ' Georges Méliès' ],
plot: 'A bat flies into an ancient castle and transforms itself into Mephistopheles himself. Producing a cauldron, Mephistopheles conjures up a young girl and various supernatural creatures, one ...',
fullplot: 'A bat flies into an ancient castle and transforms itself into Mephistopheles himself. Producing a cauldron, Mephistopheles conjures up a young girl and various supernatural creatures, one of which brandishes a crucifix in an effort to force the devil-vampire to vanish.',
lastupdated: '2015-08-26 00:06:16.697000000',
type: 'movie',
directors: [ 'Georges Méliès'],
writers: [ ' Georges Méliès'],
imdb: { rating: 6.8, votes: 1135, id: 91 },
countries: [ 'France' ],
genres: [ 'Short', 'Horror' ],
tomatoes: {
viewer: { rating: 5, numReviews: 23 },
lastUpdated: ISODate("2015-06-02T19:48:08.000Z") },
num_mflix_comments: 1,
comments: [{
  name: 'Oscar Sanchez',
  email: 'oscar_sanchez@fakegmail.com',
  movie_id: ObjectId("573a1390f29313caabcd418c"),
  text: 'Non repellat atque in ipsa accusantium. Assumenda modi magni quis.\n' +
    'Recusandae recusandae dicta repellat ad reprehenderit mollitia quam. Itaque voluptate asperiores quia alias.',
  date: ISODate("2017-02-11T07:00:52.000Z")
} ] }

```

1. Đếm tổng số các document movies.
2. Xuất các document movies theo năm, tính tổng số film trong mỗi năm
3. Xuất các document movies theo năm, tính tổng số film trong mỗi năm, sau đó sắp xếp tăng.
4. Xuất các document movies theo năm, sau đó sắp xếp theo thứ tự giảm dần dựa trên số lượng.
5. Xuất các document movies theo số lượng film mỗi đạo diễn có được,
6. Xuất các document movies theo số lượng film từng year, type, title. Sau đó sắp xếp giảm dần theo số lượng đếm được.
7. Liệt kê danh sách các đạo diễn có tham gia từ 30 bộ phim trở lên. Thông tin bao gồm: Tên đạo diễn (director) và số bộ phim.
8. Tìm tất cả các đạo diễn có tham gia đạo diễn nhiều bộ phim nhất
9. Liệt kê tựa phim (title) theo từng đạo diễn. Thông tin bao gồm: tên đạo diễn (director) và danh sách tựa phim
10. Thống kê số bộ phim đã phát hành theo từng năm, sắp xếp giảm dần theo năm
11. Tìm năm phát hành nhiều bộ phim nhất.
12. Liệt kê danh sách các tựa phim (title) theo từng quốc gia. Thông tin bao gồm: tên quốc gia và danh sách tựa phim
13. Đếm số bộ phim theo từng quốc gia, sắp xếp giảm dần theo số bộ phim. Thông tin bao gồm: Tên quốc gia và số bộ phim

14. Tìm những tựa phim (title) phát hành trong tháng 03 năm 2016

15. Liệt kê những tựa phim (title) do diễn viên “Frank Powell” hoặc “Charles Wellesley” đóng

16. Tìm những quốc gia phát hành nhiều bộ phim nhất

FIT-SE