

CSC 369: Distributed Computing

Alex Dekhtyar

April 20

Day 7: MongoDB Aggregation Pipeline, Part II



Housekeeping: plan for the week

April 20

- Lecture: joins
- Lab: practice
- Assignment: Lab 3 **(Lab 3-2 by Tuesday)**

Grading:

Lab 1: not certain

Lab 2: end of the week

April 22

- Lecture: catch-all on MongoDB, Mongo Python connectivity
- Lab: practice
- Assignment: Lab 3

April 24

- Lecture: MongoDB quiz (?)
- Lab: practice
- Assignment: Lab 3 and Lab 4 (build a MongoDB-supported application)

April 27 (Monday)

- MongoDB test
- Flip lecture/lab

Recap of Last Week

Data Manipulation operations

Filtering

Given a condition - keep only objects that satisfy it

Projection Transformation

Modify the contents of its object based solely on what's in the object itself

Grouping

Break collection into groups, each representing objects with same values of some keys

Aggregation

Compute an aggregate value over a set of objects

Join

Combine objects from two different collections based on matches in values of some keys

Sort

Return objects in a specific order

Data Manipulation Operations

**Ungrouping
Unwinding**

Opposite of grouping - build an object for each element of an array

Limit

Return a specific number of documents

Skip

Return documents after skipping a specified number

Sample

Return a random sample of documents

\$operation

Filtering

\$match

\$unwind

Unwinding

Projection

\$project

\$limit

Limit

Aggregation

\$group

\$skip

Skip

Grouping

\$sample

Sample

Sort

\$sort

Join

\$lookup

\$operation

Filtering

\$match

\$redact

Projection

\$project

\$set

\$unset

\$addField

\$replaceRoot

Aggregation

Grouping

\$group

\$bucket

\$bucketAuto

Sort

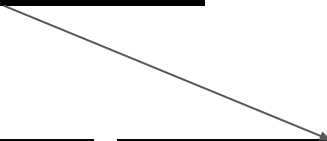
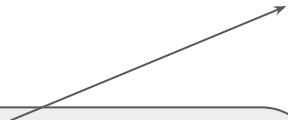
\$sort

\$sortByCount

Join

\$lookup

\$graphLookup



\$operation

Filtering

\$match

\$redact

Projection

\$project

\$set

\$unset

\$addField

\$replaceRoot

Aggregation

Grouping

\$group

\$bucket

\$bucketAuto

Sort

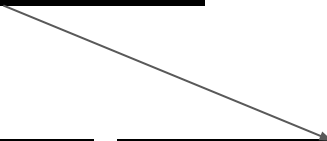
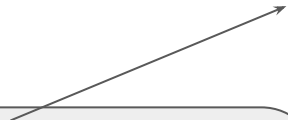
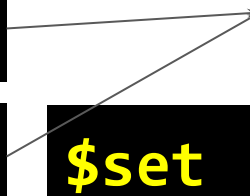
\$sort

\$sortByCount

Join

\$lookup

\$graphLookup



\$operation

Join

\$lookup

Faceted Filter

\$facet

Output Control

\$out

\$merge

Joins

Combine information from multiple data sources

In Relational Algebra: cartesian product followed by selection

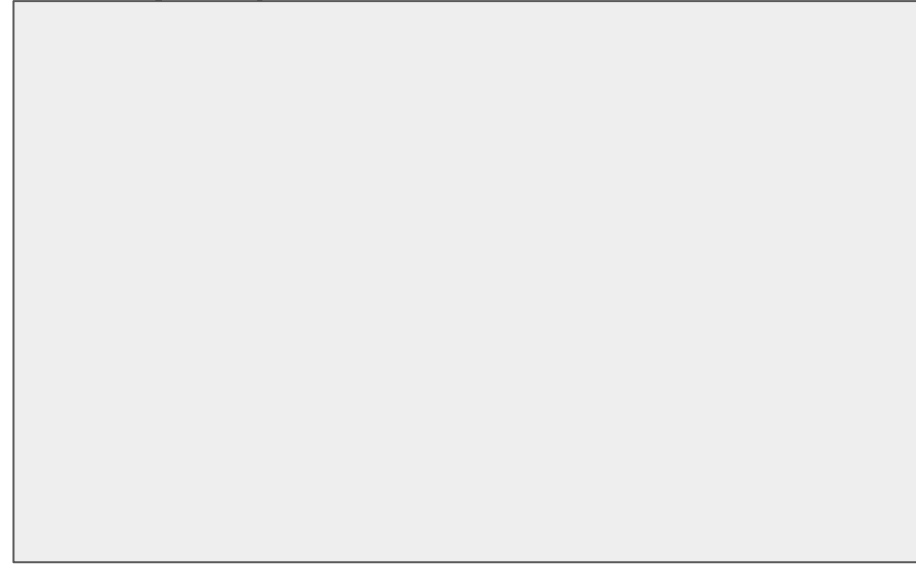
Joins (Relational Algebra Example)

Orders		
Order	Product	Price
1234	1001	\$4.99
1234	2031	\$20.00
3403	2031	\$20.00
3311	1100	\$12.00

Products		
Product	Name	Rating
1001	Hand Sanitizer	4.5
1100	Gloves	4
2031	Masks	3
4011	Teddy Bear	5

Joins (Relational Algebra Example)

Orders		
Order	Product	Price
1234	1001	\$4.99
1234	2031	\$20.00
3403	2031	\$20.00
3311	1100	\$12.00



What product was purchased?

“1001” is NOT a good answer

Joins (Relational Algebra Example)

Orders		
Order	Product	Price
1234	1001	\$4.99
1234	2031	\$20.00
3403	2031	\$20.00
3311	1100	\$12.00

Products		
Product	Name	Rating
1001	Hand Sanitizer	4.5
1100	Gloves	4
2031	Masks	3
4011	Teddy Bear	5



Joins (Relational Algebra Example)

Orders		
Order	Product	Price
1234	1001	\$4.99
1234	2031	\$20.00
3403	2031	\$20.00
3311	1100	\$12.00

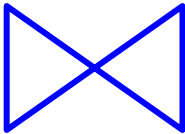
Products		
Product	Name	Rating
1001	Hand Sanitizer	4.5
1100	Gloves	4
2031	Masks	3
4011	Teddy Bear	5



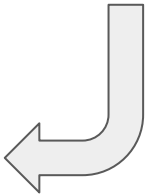
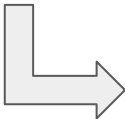
Joins (Relational Algebra Example)

Orders		
Order	Product	Price
1234	1001	\$4.99
1234	2031	\$20.00
3403	2031	\$20.00
3311	1100	\$12.00

Products		
Product	Name	Rating
1001	Hand Sanitizer	4.5
1100	Gloves	4
2031	Masks	3
4011	Teddy Bear	5

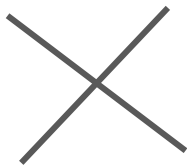


Order	Product	Price	Name	Rating
1234	1001	\$4.99	Hand Sainitizer	4.5
1234	2031	\$20.00	Masks	3
3403	2031	\$20.00	Masks	3
...



Joins (Relational Algebra Example)

Orders		
Order	Product	Price
1234	1001	\$4.99
1234	2031	\$20.00
3403	2031	\$20.00
3311	1100	\$12.00

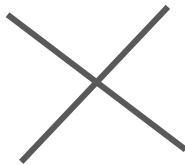


Products		
Product	Name	Rating
1001	Hand Sanitizer	4.5
1100	Gloves	4
2031	Masks	3
4011	Teddy Bear	5

Order	O.Product	Price	P.Product	Name	Rating
1234	1001	\$4.99	1001	Hand Sanitizer	4.5
1234	1001	\$4.99	1100	Gloves	4
1234	1001	\$4.99	2031	Masks	3
1234	1001	\$4.99	4011	Teddy Bear	5
1234	2031	\$20.00	1001	Hand Sanitizer	4.5

Joins (Relational Algebra Example)

Orders		
Order	Product	Price
1234	1001	\$4.99
1234	2031	\$20.00
3403	2031	\$20.00
3311	1100	\$12.00

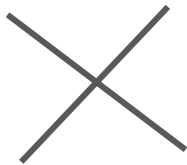


Products		
Product	Name	Rating
1001	Hand Sanitizer	4.5
1100	Gloves	4
2031	Masks	3
4011	Teddy Bear	5

Order	O.Product	Price	P.Product	Name	Rating
1234	1001	\$4.99	1001	Hand Sanitizer	4.5
1234	1001	\$4.99	1100	Gloves	4
1234	1001	\$4.99	2031	Masks	3
1234	1001	\$4.99	4011	Teddy Bear	5
1234	2031	\$20.00	1001	Hand Sanitizer	4.5

Joins (Relational Algebra Example)

Orders		
Order	Product	Price
1234	1001	\$4.99
1234	2031	\$20.00
3403	2031	\$20.00
3311	1100	\$12.00



Products		
Product	Name	Rating
1001	Hand Sanitizer	4.5
1100	Gloves	4
2031	Masks	3
4011	Teddy Bear	5

Order	O.Product	Price	P.Product	Name	Rating
1234	1001	\$4.99	1001	Hand Sanitizer	4.5
1234	1001	\$4.99	1100	Gloves	4
1234	1001	\$4.99	2031	Masks	3
1234	1001	\$4.99	4011	Teddy Bear	5
1234	2031	\$20.00	1001	Hand Sanitizer	4.5

Joins (Relational Algebra Example)

Orders		
Order	Product	Price
1234	1001	\$4.99
1234	2031	\$20.00
3403	2031	\$20.00
3311	1100	\$12.00

Cartesian product

Selection

Products		
Product	Name	Rating
1001	Hand Sanitizer	4.5
1100	Gloves	4
2031	Masks	3
4011	Teddy Bear	5

Order	O.Product	Price	P.Product	Name	Rating
1234	1001	\$4.99	1001	Hand Sanitizer	4.5
1234	1001	\$4.99	1100	Gloves	4
1234	1001	\$4.99	2031	Masks	3
1234	1001	\$4.99	4011	Teddy Bear	5
1234	2031	\$20.00	1001	Hand Sanitizer	4.5

Joins (Relational Algebra Example)

Orders		
Order	Product	Price
1234	1001	\$4.99
1234	2031	\$20.00
3403	2031	\$20.00
3311	1100	\$12.00

Cartesian product

Selection

Products		
Product	Name	Rating
1001	Hand Sanitizer	4.5
1100	Gloves	4
2031	Masks	3
4011	Teddy Bear	5

Order	O.Product	Price	P.Product	Name	Rating
1234	1001	\$4.99	1001	Hand Sanitizer	4.5
1234	1001	\$4.99		Gloves	4
1234	1001	\$4.99		Masks	3
1234	1001	\$4.99		Teddy Bear	5
1234	2031	\$20.00		Hand Sanitizer	4.5

Joins (Relational Algebra Example)

Orders		
Order	Product	Price
1234	1001	\$4.99
1234	2031	\$20.00
3403	2031	\$20.00
3311	1100	\$12.00

Cartesian product

Selection

Projection

Products		
Product	Name	Rating
1001	Hand Sanitizer	4.5
1100	Gloves	4
2031	Masks	3
4011	Teddy Bear	5

Order	O.Product	Price	Name	Rating
1234	1001	\$4.99	Hand Sanitizer	4.5
1234	1001	\$4.99	Gloves	4
...
1234	1001	\$4.99	Teddy Bear	5
1234	2031	\$20.00	Hand Sanitizer	4.5

Inner Joins vs Outer Joins

Orders		
Order	Product	Price
1234	1001	\$4.99
1234	2031	\$20.00
3403	2031	\$20.00
3311	1100	\$12.00

Products		
Product	Name	Rating
1001	Hand Sanitizer	4.5
1100	Gloves	4
2031	Masks	3
4011	Teddy Bear	5

Order	Product	Price	Name	Rating
1234	1001	\$4.99	Hand Sanitizer	4.5

For each product ordered report the name of the product and the rating.

...

Inner Joins vs Outer Joins

Orders			Products		
Order	Product				Rating
1234	1001	\$4.99	1001	Hand Sanitizer	4.5
1234	2031	\$20.00	1100	Gloves	4
3403	2031	\$20.00	2031	Masks	3
3311	1100	\$12.00	4011	Teddy Bear	5

For each product, report the details of all orders.

Order	Price	Product	Name	Rating
1234	\$4.99	1001	Hand Sanitizer	4.5
3311	\$12.00	1100	Gloves	4
1234	\$20.00	2031	Masks	3
3403	\$20.00	2031	Masks	3

Inner Joins vs Outer Joins

Orders			Products		
Order	Product				Rating
1234	1001	\$4.99	1001	Hand Sanitizer	4.5
1234	2031	\$20.00	1100	Gloves	4
3403	2031	\$20.00	2031	Masks	3
3311	1100	\$12.00	4011	Teddy Bear	5

For each product, report the details of all orders.

Order	Price	Product	Name	Rating
1234	\$4.99	1001	Hand Sanitizer	4.5
3311	\$12.00	1100	Gloves	4
1234	\$20.00	2031	Masks	3
3403	\$20.00	2031	Masks	3

Inner Joins vs Outer Joins

Orders			Products				
Order	Product		For each product, report the details of all orders. Keep all products in output				Rating
1234	1001	\$4.99		1001	Hand Sanitizer	4.5	
1234	2031	\$20.00	Right outer join	1100	Gloves	4	
3403	2031	\$20.00		2031	Masks	3	
3311	1100	\$12.00		4011	Teddy Bear	5	

Order	Price	Product	Name	Rating
1234	\$4.99	1001	Hand Sanitizer	4.5
3311	\$12.00	1100	Gloves	4
1234	\$20.00	2031	Masks	3
3403	\$20.00	2031	Masks	3
NULL	NULL	4011	Teddy Bear	5

Joins in MongoDB

```
{ $lookup: {  
    from: <collection>,  
    localField: <field>,  
    foreignField: <field>,  
    as: <arrayField>  
}  
}
```

Joins in MongoDB

```
{ $lookup: {  
    from: <collection>,  
    localField: <field>,  
    foreignField: <field>,  
    as: <arrayField>  
  }  
}
```

Left outer join

Joins in MongoDB

Left outer join

```
{ $lookup: {  
    from: <collection>,  
    localField: <field>,  
    foreignField: <field>,  
    as: <arrayField>  
  }  
}
```

Join condition: *localField == collection.foreignField*

Joins in MongoDB

Left outer join

```
{ $lookup: {  
    from: <collection>,  
    localField: <field>,  
    foreignField: <field>,  
    as: <arrayField>  
  }  
}
```

Join condition: *localField == collection.foreignField*

Matched documents are imported into *<arrayField>*

Joins in MongoDB

Left outer join

EXAMPLE

```
{ $lookup: {  
    from: <collection>,  
    localField: <field>,  
    foreignField: <field>,  
    as: <arrayField>  
  }  
}
```

Join condition: *localField == collection.foreignField*

Matched documents are imported into *<arrayField>*

```
db.spring.aggregate(  
  {$project: {_id:0, name:1, department:1}},  
  {$lookup: { from: "departments",  
              localField:"department",  
              foreignField: "department",  
              as: "deptInfo"  
            }  
  },  
  {$unwind: "$deptInfo"}  
)
```



```
db.departments.aggregate(  
  {$lookup: {from: "spring",  
             localField: "department",  
             foreignField: "department",  
             as: "faculty"  
            }  
}  
)
```

Facets: set up multiple independent pipelines

```
{ $facet: { key1: [<stage1>, <stage2>, ..., <stageK1>] ,  
            key2: [<stage1>, <stage2>, ..., <stageK2>] ,  
            ...  
            keyN:  [<stage1>, <stage2>, ..., <stageKn>]  
          }  
}
```