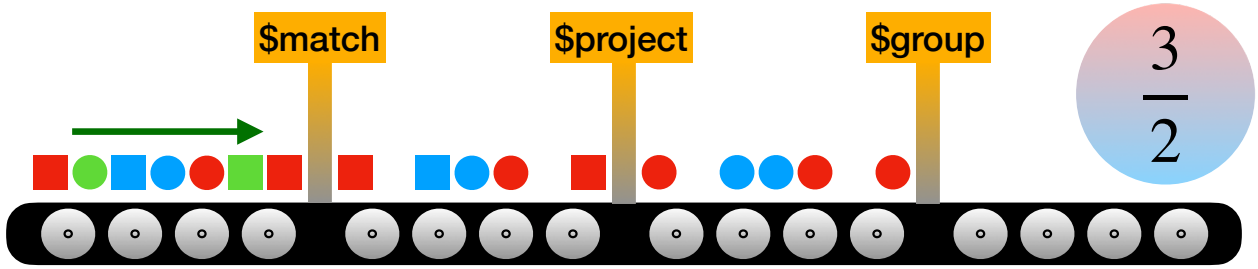
MongoDB Aggregation Pipelines

COSC 061 - Fall 2022 - Dartmouth College

Pipelines

Think “Assembly line”

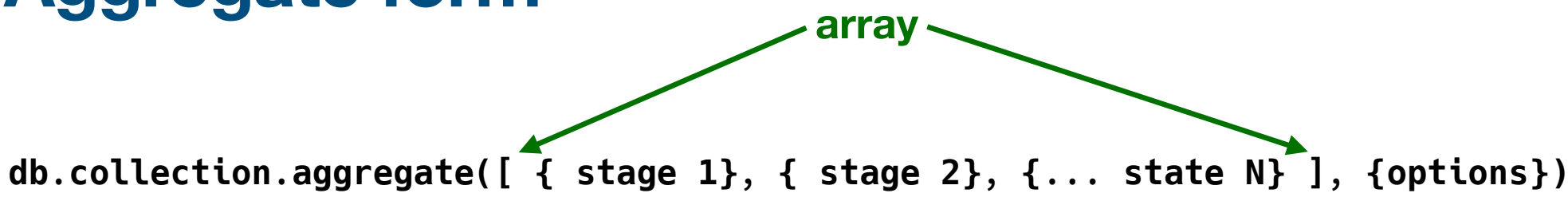


$match filters out the green shapes

$project turns squares into circles

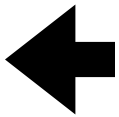
$group finds the ratio of red to blue circles (documents) and produces a single document containing that ratio.

Aggregate form

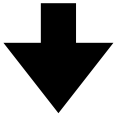


SQL Example

```
SELECT make, COUNT(*) as numModels
FROM dirtBikes
GROUP BY model;
```



make	model	cc	priceNew
Yamaha	TTR50	50	1500
Yamaha	TTR110	110	2100
Honda	CRF50F	50	1500
Honda	CRF125	125	2999
Honda	CRF125BW	125	3399
Kawasaki	KX65	65	3700



make	numModels
Yamaha	2
Honda	3
Kawasaki	1

MongoDB Aggregate

dirtBikes

	_id ObjectId	make String	model String	cc Int32	priceNew Int32
1	ObjectId('63643664c4a499079…	"Yamama"	"TTR50"	50	1500
2	ObjectId('63643664c4a499079…	"Yamama"	"TTR110"	110	2100
3	ObjectId('63643664c4a499079…	"Honda"	"CRF50"	50	1500
4	ObjectId('63643664c4a499079…	"Honda"	"CRF125"	125	3000
5	ObjectId('63643664c4a499079…	"Honda"	"CRF125BW"	125	3400
6	ObjectId('63643664c4a499079…	"Kawasaki"	"KX65"	65	3700

```
db.dirtBikes.aggregate([{
    $group: {
        "_id": "$make",
        "num_models": {
            $sum: 1
        }
    }
}])
```

```
{ _id: 'Yamama', num_models: 2 }
{ _id: 'Honda', num_models: 3 }
{ _id: 'Kawasaki', num_models: 1 }
```

MongoDB Aggregate

dirtBikes

	_id ObjectId	make String	model String	cc Int32	priceNew Int32
1	ObjectId('63643664c4a499079…	"Yamama"	"TTR50"	50	1500
2	ObjectId('63643664c4a499079…	"Yamama"	"TTR110"	110	2100
3	ObjectId('63643664c4a499079…	"Honda"	"CRF50"	50	1500
4	ObjectId('63643664c4a499079…	"Honda"	"CRF125"	125	3000
5	ObjectId('63643664c4a499079…	"Honda"	"CRF125BW"	125	3400
6	ObjectId('63643664c4a499079…	"Kawasaki"	"KX65"	65	3700

```
db.dirtBikes.aggregate([{
    $match: {
         "priceNew" : {$gt : 2700 }
        }
    },{
    $group: {
        "_id": "$make",
        "num_models": {
            $sum: 1
        },
        "avgCost": {
            $avg: "$priceNew"
        }
    }
}])
```

```
{ _id: 'Honda', num_models: 2, avgCost: 3200 }
{ _id: 'Kawasaki', num_models: 1, avgCost: 3700 }
```

stage	action	Doc mapping
$project	selects values from the documents, even deep inside.	1:1
$match	filter	n:1
$group	aggregation, like sum, COUNT	n:1
$sort	sorting	1:1
$skip	skips results	n:1
$limit	limit the results	n:1
$unwind	flatten or denormalize the data	1:n
$out	redirect output to a collection instead of a cursor	1:1
$redact	limits documemts based on access permission info	m:n, m >= n
$geonear	limits documemts based on location info	m:n, m >= n