Нереляционные базы данных

Лабораторная работа 5

MapReduce

1. Создайте коллекцию с 5-10 документами с 3-5 полями, содержащими числа, строки и массивы. Для каждого поля (или комбинации полей) выполните произвольных MapReduce и выведите результат на экран. Результат одного из MapReduce сохраните в новую коллекцию.

```
> db.humanFruits.insert({name: "Alex", age: 15, fruits:["apple", "banana"]})
WriteResult({ "nInserted" : 1 })
> db.humanFruits.insert({name: "Misha", age: 13, fruits:["apple", "banana", "orange"]})
WriteResult({ "nInserted" : 1 })
> db.humanFruits.insert({name: "Kirill", age: 13, fruits:[ "banana", "orange"]})
WriteResult({ "nInserted" : 1 })
> db.humanFruits.insert({name: "Kirill", age: 12, fruits:[ "orange"]})
WriteResult({ "nInserted" : 1 })
```

```
> var map = function(){ var key = {name: this.name, age:this.age}; var count = {val: this.fruits}; emit(key, {count: count.val.length}) };
```

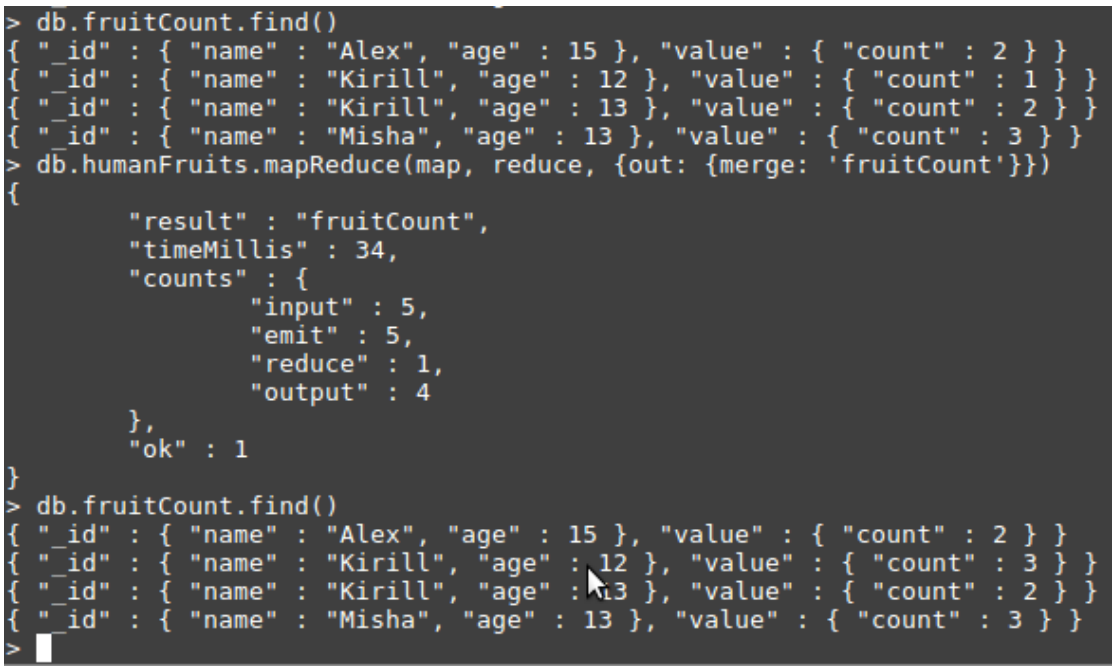
```
> db.humanFruits.find()
{ "_id" : ObjectId("5f7885e0252e02743d540d5d"), "name" : "Alex", "age" : 15, "fruits" : [ "apple", "banana" ] }
{ "_id" : ObjectId("5f788602252e02743d540d5e"), "name" : "Misha", "age" : 13, "fruits" : [ "apple", "banana", "orange" ] }
{ "_id" : ObjectId("5f788612252e02743d540d5f"), "name" : "Kirill", "age" : 13, "fruits" : [ "banana", "orange" ] }
{ "_id" : ObjectId("5f78861e252e02743d540d60"), "name" : "Kirill", "age" : 12, "fruits" : [ "orange" ] }
> var reduce = function(key, values) {
...  var sum = 0;
...  values.forEach(function(value) {
...  sum += value['count'];
...  });
...  return {count: sum};
... };
```

```
} at src/mongo/shell/collection.js:1224
> db.humanFruits.mapReduce(map, reduce, {out: {inline:1}})
{
        "results" : [
                {
                        "_id" : {
                                "name" : "Alex",
                                "age" : 15
                        },
                        "value" : {
                                "count" : 2
                        }
                },
                {
                        "_id" : {
                                "name" : "Kirill",
                                "age" : 12
                        },
                        "value" : {
                                "count" : 1
                        }
                },
                {
                        "_id" : {
                                "name" : "Kirill",
                                "age" : 13
                        },
                        "value" : {
                                "count" : 2
                        }
                },
                {
                        "_id" : {
                                "name" : "Misha",
                                "age" : 13
                        },
                        "value" : {
                                "count" : 3
                        }
```

```
> db.humanFruits.mapReduce(map, reduce, {out: 'fruitCount'})
{
        "result" : "fruitCount",
        "timeMillis" : 125,
        "counts" : {
                "input" : 4,
                "emit" : 4,
                "reduce" : 0,
                "output" : 4
        },
        "ok" : 1
}
> db.fruitCount.find()
{ "_id" : { "name" : "Alex", "age" : 15 }, "value" : { "count" : 2 } }
{ "_id" : { "name" : "Kirill", "age" : 12 }, "value" : { "count" : 1 } }
{ "_id" : { "name" : "Kirill", "age" : 13 }, "value" : { "count" : 2 } }
{ "_id" : { "name" : "Misha", "age" : 13 }, "value" : { "count" : 3 } }
>
```

2. При наличии интернета изучите 1-3 дополнительных значения поля out, продемонстрируйте их работу на произвольных коллекциях из задания 1.

```
> db.humanFruits.mapReduce(map, reduce, {out: {replace: 'unicorns'}})
{
        "result" : "unicorns",
        "timeMillis" : 16,
        "counts" : {
                "input" : 5,
                "emit" : 5,
                "reduce" : 1,
                "output" : 4
        },
        "ok" : 1
}
> db.unicorns.find()
{ "_id" : { "name" : "Alex", "age" : 15 }, "value" : { "count" : 2 } }
{ "_id" : { "name" : "Kirill", "age" : 12 }, "value" : { "count" : 3 } }
{ "_id" : { "name" : "Kirill", "age" : 13 }, "value" : { "count" : 2 } }
{ "_id" : { "name" : "Misha", "age" : 13 }, "value" : { "count" : 3 } }
>
```



3. Сделайте вывод о плюсах и минусах MapReduce.

+ Можно распараллелить код

+ Можно описывать обработку обычным кодом

- Нужно хотя-бы в общих чертах знать язык