BIG DATA LAB

P.SESHA HARSHINI
1NT19IS107
C1 BATCH
31-05-2022

1 Problem Statement & Dataset

1.Create a collection named "Employee" under the "EmployeeDB" database with each document in the format shown below Table 1.

Table 1: Document Format

Name	Age	Salary in INR	Designation	Role
{Firstname, middlename, lastname}	25-40	20000 - 75000	Employee Designation	[Manager, "Team Lead", "Software Developer", "Tester", "UI Designer"]
String BSON Object	int	Number	String	String Array

use harshiniEmployeeDB
db.createCollection("Employee")

```
> use harshiniEmployeeDB
switched to db harshiniEmployeeDB
> db.createCollection("Employee")
{ "ok" : 1 }
> show collections
Employee
```

Queries

1. Populate the database with atleast 15 documents

```
db.Employee.insertMany([{firstname:"P",middlename:"Sesha",lastname:"Harshini",age:25,salary:50000,designation:"Lead",role:"software developer"},{firstname:"Y",middlename:"Mahitha",lastname:"Reddy",age:28,salary:40000,designation:"HR",role:"Tester"},{firstname:"P",middlename:"Snehitha",lastname:"Chowdary",age:30,salary:45000,designation:"Associate",role:"UI Designer"} ])
```

```
db.Employee.insertMany([{firstname:"Mahesh",middlename:"Kumar",lastname:"Reddy",age:30,salary:35000,designation:"scientist",role:"software developer"},{firstname:"Rajesh",middlename:"Reddy",lastname:"C",age:35,salary:60000,designation:"scientist",role:"Manager"},{firstname:"T",middlename:"Nikitha",lastname:"Reddy",age:40,salary:45000,designation:"Recruiter",role:"tester"} ])
```

```
db.Employee.insertMany([{firstname:"Allu",middlename:"Indu",lastname:"Chowdary",age:37,salary:70000,designation:"HR",role:"Team Lead"},{firstname:"Parimi",middlename:"Teja",lastname:"Sree",age:25,salary:37000,designation:"HR",role:"Tester"},{firstname:"G",middlename:"Krishna",lastname:"Sai",age:38,salary:55650,designation:"Analyst",role:"Team Lead"} ])
```

```
db.Employee.insertMany([{firstname:"S",middlename:"L",lastname:"laisha",age:27,salary:50000,designation:"Associate",role:"UI Designer"},{firstname:"P",middlename:"Chaitanya",lastname:"Chowdary",age:30,salary:45000,designation:"Guide",role:"software developer"},{firstname:"Potula",middlename:"Divya",lastname:"Sree",age:40,salary:62300,designation:"Leader",role:"Manager"} ])
```

```
db.Employee.insertMany([{firstname:"S",middlename:"R",lastname:"Sahana",age:27,salary:32000,designation:"Analyst",role:"software developer"},{firstname:"B",middlename:"Deeksha",lastname:"Banoth",age:33,salary:31500,designation:"Guide",role:"Tester"},{firstname:"P",middlename:"Jayanth",lastname:"Sai",age:30,salary:30000,designation:"Analyst",role:"Team Lead"} ])
```

```
db.Employee.find()
```

```
> db.Employee.find()
{ "_id" : ObjectId("629596b4d4d91df5c5e3f03c"), "firstname" : "P", "middlename" : "Sesha", "lastname" : "Harshini", "age" : 25, "salary" : 500
00, "designation" : "Lead", "role" : "software developer" }
{ "_id" : ObjectId("629596b4d4d91df5c5e3f03d"), "firstname" : "Y", "middlename" : "Mahitha", "lastname" : "Reddy", "age" : 28, "salary" : 4000
0, "designation" : "HR", "role" : "Tester" }
{ "_id" : ObjectId("629596b4d4d91df5c5e3f03e"), "firstname" : "P", "middlename" : "Snehitha", "lastname" : "Chowdary", "age" : 30, "salary" :
45000, "designation" : "Associate", "role" : "UI Designer" }
{ "_id" : ObjectId("629597a8d4d91df5c5e3f03f"), "firstname" : "Mahesh", "middlename" : "Kumar", "lastname" : "Reddy", "age" : 30, "salary" : 3
5000, "designation" : "scientist", "role" : "software developer" }
{ "_id" : ObjectId("629597a8d4d91df5c5e3f040"), "firstname" : "Rajesh", "middlename" : "Reddy", "lastname" : "C", "age" : 35, "salary" : 60000
, "designation" : "scientist", "role" : "Manager" }
{ "_id" : ObjectId("629597a8d4d91df5c5e3f041"), "firstname" : "T", "middlename" : "Nikitha", "lastname" : "Reddy", "age" : 40, "salary" : 4500
0, "designation" : "Recruiter", "role" : "tester" }
{ "_id" : ObjectId("6295989dd4d91df5c5e3f042"), "firstname" : "Allu", "middlename" : "Indu", "lastname" : "Chowdary", "age" : 37, "salary" : 7
0000, "designation" : "HR", "role" : "Team Lead" }
{ "_id" : ObjectId("6295989dd4d91df5c5e3f043"), "firstname" : "Parimi", "middlename" : "Teja", "lastname" : "Sree", "age" : 25, "salary" : 370
00, "designation" : "HR", "role" : "Tester" }
{ "_id" : ObjectId("6295989dd4d91df5c5e3f044"), "firstname" : "G", "middlename" : "Krishna", "lastname" : "Sai", "age" : 38, "salary" : 55650,
 "designation" : "Analyst", "role" : "Team Lead" }
{ "_id" : ObjectId("629599dcd4d91df5c5e3f045"), "firstname" : "S", "middlename" : "L", "lastname" : "laisha", "age" : 27, "salary" : 50000, "d
esignation" : "Associate", "role" : "UI Designer" }
{ "_id" : ObjectId("629599dcd4d91df5c5e3f046"), "firstname" : "P", "middlename" : "Chaitanya", "lastname" : "Chowdary", "age" : 30, "salary" :
 45000, "designation" : "Guide", "role" : "software developer" }
{ "_id" : ObjectId("629599dcd4d91df5c5e3f047"), "firstname" : "Potula", "middlename" : "Divya", "lastname" : "Sree", "age" : 40, "salary" : 62
300, "designation" : "Leader", "role" : "Manager" }
{ "_id" : ObjectId("62959b3bd4d91df5c5e3f048"), "firstname" : "S", "middlename" : "R", "lastname" : "Sahana", "age" : 27, "salary" : 32000, "d
esignation" : "Analyst", "role" : "software developer" }
{ "_id" : ObjectId("62959b3bd4d91df5c5e3f049"), "firstname" : "B", "middlename" : "Deeksha", "lastname" : "Banoth", "age" : 33, "salary" : 315
00, "designation" : "Guide", "role" : "Tester" }
{ "_id" : ObjectId("62959b3bd4d91df5c5e3f04a"), "firstname" : "P", "middlename" : "Jayanth", "lastname" : "Sai", "age" : 30, "salary" : 30000,
 "designation" : "Analyst", "role" : "Team Lead" }
>
```

2. List all the records having salary in the range of 20000 - 35000(Exclusive)

db.Employee.find({$and:[{salary:{$gt:20000}},{salary:{$lt:35000}}]})

```
> db.Employee.find({$and:[{salary:{$gt:20000}},{salary:{$lt:35000}}]})
{ "_id" : ObjectId("62959b3bd4d91df5c5e3f048"), "firstname" : "S", "middlename" : "R", "lastname" : "Sahana", "age" : 27, "salary" : 32000, "d
esignation" : "Analyst", "role" : "software developer" }
{ "_id" : ObjectId("62959b3bd4d91df5c5e3f049"), "firstname" : "B", "middlename" : "Deeksha", "lastname" : "Banoth", "age" : 33, "salary" : 315
00, "designation" : "Guide", "role" : "Tester" }
{ "_id" : ObjectId("62959b3bd4d91df5c5e3f04a"), "firstname" : "P", "middlename" : "Jayanth", "lastname" : "Sai", "age" : 30, "salary" : 30000,
 "designation" : "Analyst", "role" : "Team Lead" }
>
```

3. List all the Employee whose Middle name is "Kumar"

db.Employee.find({middlename:"Kumar"})

```
> db.Employee.find({middlename:"Kumar"})
{ "_id" : ObjectId("629597a8d4d91df5c5e3f03f"), "firstname" : "Mahesh", "middlename" : "Kumar", "lastname" : "Reddy", "age" : 30, "salary" : 3
5000, "designation" : "scientist", "role" : "software developer" }
```

4. Count the number of Employees who has a role "Manager" in the Role field.

db.Employee.count({role:"Manager"})

```
> db.Employee.count({role:"Manager"})
2
```

5. Find out all the documents who have age < 35 and salary in the range of 30000-35000.

db.Employee.find({$and:[{salary:{$gt:30000}},{salary:{$lt:35000}},{age:{$lt:35}}]})

```
> db.Employee.find({$and:[{salary:{$gt:30000}},{salary:{$lt:35000}},{age:{$lt:35}}]})
{ "_id" : ObjectId("62959b3bd4d91df5c5e3f048"), "firstname" : "S", "middlename" : "R", "lastname" : "Sahana", "age" : 27, "salary" : 32000, "d
esignation" : "Analyst", "role" : "software developer" }
{ "_id" : ObjectId("62959b3bd4d91df5c5e3f049"), "firstname" : "B", "middlename" : "Deeksha", "lastname" : "Banoth", "age" : 33, "salary" : 315
00, "designation" : "Guide", "role" : "Tester" }
```

6. Delete an Employee whose "Firstname" is "Rajesh" and having the designation as "Scientist".

db.Employee.deleteOne({$and:[{firstname:"Rajesh"},{designation:"scientist"}]})

db.Employee.find()

```
> db.Employee.deleteOne({$and:[{firstname:"Rajesh"},{designation:"scientist"}]})
{ "acknowledged" : true, "deletedCount" : 1 }
> db.Employee.find()
{ "_id" : ObjectId("629596b4d4d91df5c5e3f03c"), "firstname" : "P", "middlename" : "Sesha", "lastname" : "Harshini", "age" : 25, "salary" : 500
00, "designation" : "Lead", "role" : "software developer" }
{ "_id" : ObjectId("629596b4d4d91df5c5e3f03d"), "firstname" : "Y", "middlename" : "Mahitha", "lastname" : "Reddy", "age" : 28, "salary" : 4000
0, "designation" : "HR", "role" : "Tester" }
{ "_id" : ObjectId("629596b4d4d91df5c5e3f03e"), "firstname" : "P", "middlename" : "Snehitha", "lastname" : "Chowdary", "age" : 30, "salary" :
45000, "designation" : "Associate", "role" : "UI Designer" }
{ "_id" : ObjectId("629597a8d4d91df5c5e3f03f"), "firstname" : "Mahesh", "middlename" : "Kumar", "lastname" : "Reddy", "age" : 30, "salary" : 3
5000, "designation" : "scientist", "role" : "software developer" }
{ "_id" : ObjectId("629597a8d4d91df5c5e3f041"), "firstname" : "T", "middlename" : "Nikitha", "lastname" : "Reddy", "age" : 40, "salary" : 4500
0, "designation" : "Recruiter", "role" : "tester" }
{ "_id" : ObjectId("6295989dd4d91df5c5e3f042"), "firstname" : "Allu", "middlename" : "Indu", "lastname" : "Chowdary", "age" : 37, "salary" : 7
0000, "designation" : "HR", "role" : "Team Lead" }
{ "_id" : ObjectId("6295989dd4d91df5c5e3f043"), "firstname" : "Parimi", "middlename" : "Teja", "lastname" : "Sree", "age" : 25, "salary" : 370
00, "designation" : "HR", "role" : "Tester" }
{ "_id" : ObjectId("6295989dd4d91df5c5e3f044"), "firstname" : "G", "middlename" : "Krishna", "lastname" : "Sai", "age" : 38, "salary" : 55650,
 "designation" : "Analyst", "role" : "Team Lead" }
{ "_id" : ObjectId("629599dcd4d91df5c5e3f045"), "firstname" : "S", "middlename" : "L", "lastname" : "laisha", "age" : 27, "salary" : 50000, "d
esignation" : "Associate", "role" : "UI Designer" }
{ "_id" : ObjectId("629599dcd4d91df5c5e3f046"), "firstname" : "P", "middlename" : "Chaitanya", "lastname" : "Chowdary", "age" : 30, "salary" :
 45000, "designation" : "Guide", "role" : "software developer" }
{ "_id" : ObjectId("629599dcd4d91df5c5e3f047"), "firstname" : "Potula", "middlename" : "Divya", "lastname" : "Sree", "age" : 40, "salary" : 62
300, "designation" : "Leader", "role" : "Manager" }
{ "_id" : ObjectId("62959b3bd4d91df5c5e3f048"), "firstname" : "S", "middlename" : "R", "lastname" : "Sahana", "age" : 27, "salary" : 32000, "d
esignation" : "Analyst", "role" : "software developer" }
{ "_id" : ObjectId("62959b3bd4d91df5c5e3f049"), "firstname" : "B", "middlename" : "Deeksha", "lastname" : "Banoth", "age" : 33, "salary" : 315
00, "designation" : "Guide", "role" : "Tester" }
{ "_id" : ObjectId("62959b3bd4d91df5c5e3f04a"), "firstname" : "P", "middlename" : "Jayanth", "lastname" : "Sai", "age" : 30, "salary" : 30000,
 "designation" : "Analyst", "role" : "Team Lead" }
```

7. Update all the Employees whose role is "Team Lead" with a salary of 55650 INR

db.Employee.updateMany({role:"Team Lead"},{$set:{salary:55650}})

> db.Employee.find()

```
> db.Employee.updateMany({role:"Team Lead"},{$set:{salary:55650}})
{ "acknowledged" : true, "matchedCount" : 3, "modifiedCount" : 2 }
> db.Employee.find()
{ "_id" : ObjectId("629596b4d4d91df5c5e3f03c"), "firstname" : "P", "middlename" : "Sesha", "lastname" : "Harshini", "age" : 25, "salary" : 500
00, "designation" : "Lead", "role" : "software developer" }
{ "_id" : ObjectId("629596b4d4d91df5c5e3f03d"), "firstname" : "Y", "middlename" : "Mahitha", "lastname" : "Reddy", "age" : 28, "salary" : 4000
0, "designation" : "HR", "role" : "Tester" }
{ "_id" : ObjectId("629596b4d4d91df5c5e3f03e"), "firstname" : "P", "middlename" : "Snehitha", "lastname" : "Chowdary", "age" : 30, "salary" :
45000, "designation" : "Associate", "role" : "UI Designer" }
{ "_id" : ObjectId("629597a8d4d91df5c5e3f03f"), "firstname" : "Mahesh", "middlename" : "Kumar", "lastname" : "Reddy", "age" : 30, "salary" : 3
5000, "designation" : "scientist", "role" : "software developer" }
{ "_id" : ObjectId("629597a8d4d91df5c5e3f041"), "firstname" : "T", "middlename" : "Nikitha", "lastname" : "Reddy", "age" : 40, "salary" : 4500
0, "designation" : "Recruiter", "role" : "tester" }
{ "_id" : ObjectId("6295989dd4d91df5c5e3f042"), "firstname" : "Allu", "middlename" : "Indu", "lastname" : "Chowdary", "age" : 37, "salary" : 5
5650, "designation" : "HR", "role" : "Team Lead" }
{ "_id" : ObjectId("6295989dd4d91df5c5e3f043"), "firstname" : "Parimi", "middlename" : "Teja", "lastname" : "Sree", "age" : 25, "salary" : 370
00, "designation" : "HR", "role" : "Tester" }
{ "_id" : ObjectId("6295989dd4d91df5c5e3f044"), "firstname" : "G", "middlename" : "Krishna", "lastname" : "Sai", "age" : 38, "salary" : 55650,
 "designation" : "Analyst", "role" : "Team Lead" }
{ "_id" : ObjectId("629599dcd4d91df5c5e3f045"), "firstname" : "S", "middlename" : "L", "lastname" : "laisha", "age" : 27, "salary" : 50000, "d
esignation" : "Associate", "role" : "UI Designer" }
{ "_id" : ObjectId("629599dcd4d91df5c5e3f046"), "firstname" : "P", "middlename" : "Chaitanya", "lastname" : "Chowdary", "age" : 30, "salary" :
 45000, "designation" : "Guide", "role" : "software developer" }
{ "_id" : ObjectId("629599dcd4d91df5c5e3f047"), "firstname" : "Potula", "middlename" : "Divya", "lastname" : "Sree", "age" : 40, "salary" : 62
300, "designation" : "Leader", "role" : "Manager" }
{ "_id" : ObjectId("62959b3bd4d91df5c5e3f048"), "firstname" : "S", "middlename" : "R", "lastname" : "Sahana", "age" : 27, "salary" : 32000, "d
esignation" : "Analyst", "role" : "software developer" }
{ "_id" : ObjectId("62959b3bd4d91df5c5e3f049"), "firstname" : "B", "middlename" : "Deeksha", "lastname" : "Banoth", "age" : 33, "salary" : 315
00, "designation" : "Guide", "role" : "Tester" }
{ "_id" : ObjectId("62959b3bd4d91df5c5e3f04a"), "firstname" : "P", "middlename" : "Jayanth", "lastname" : "Sai", "age" : 30, "salary" : 55650,
 "designation" : "Analyst", "role" : "Team Lead" }
```

8. Group all the Employees by their age(common age should be there) and calculate the average
salary obtained in the each group

```
db.Employee.aggregate([{$group:{_id:"$age",Average:{$avg:"$salary"}}}])
```

```
> db.Employee.aggregate([{$group:{_id:"$age",Average:{$avg:"$salary"}}}])
{ "_id" : 37, "Average" : 55650 }
{ "_id" : 27, "Average" : 41000 }
{ "_id" : 40, "Average" : 53650 }
{ "_id" : 30, "Average" : 45162.5 }
{ "_id" : 38, "Average" : 55650 }
{ "_id" : 25, "Average" : 43500 }
{ "_id" : 28, "Average" : 40000 }
{ "_id" : 33, "Average" : 31500 }
>
```

9. Apply the map-reduce to perform the above operation and obtain the results

```
var mapfunction=function(){emit(this.age,this.salary)}
var reducefunction=function(key,values){return Array.avg(values)}
db.Employee.mapReduce(mapfunction,reducefunction,{'out':'result'})
db.result.find()
```

```
> var mapfunction=function(){emit(this.age,this.salary)}
> var reducefunction=function(key,values){return Array.avg(values)}
> db.Employee.mapReduce(mapfunction,reducefunction,{'out':'result'})
{ "result" : "result", "ok" : 1 }
> db.result.find()
{ "_id" : 37, "value" : 55650 }
{ "_id" : 27, "value" : 41000 }
{ "_id" : 40, "value" : 53650 }
{ "_id" : 30, "value" : 45162.5 }
{ "_id" : 38, "value" : 55650 }
{ "_id" : 25, "value" : 43500 }
{ "_id" : 28, "value" : 40000 }
{ "_id" : 33, "value" : 31500 }
```