Praneeth M V L S S S
1NT19IS112
C1 -BATCH

BIG DATA

1 Problem Statement & Dataset

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

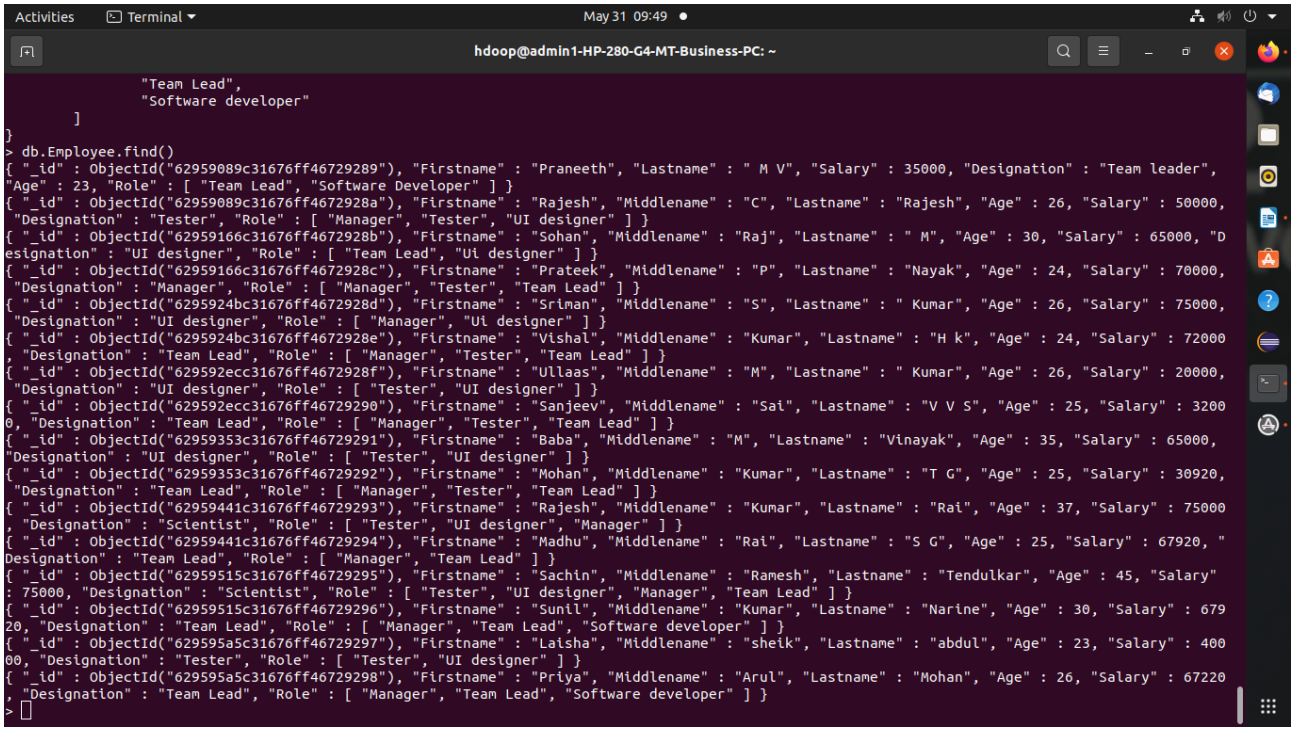
Table 1: Document Format

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

1)Populate the database with at least 15 documents

> db.Employee.find()
{ "_id" : ObjectId("62959089c31676ff46729289"), "Firstname" : "Praneeth", "Lastname" : " M V", "Salary" : 35000, "Designation" : "Team leader", "Age" : 23, "Role" : ["Team Lead", "Software Developer"] }
{ "_id" : ObjectId("62959089c31676ff4672928a"), "Firstname" : "Rajesh", "Middlename" : "C", "Lastname" : "Rajesh", "Age" : 26, "Salary" : 50000, "Designation" : "Tester", "Role" : ["Manager", "Tester", "UI designer"] }
{ "_id" : ObjectId("62959166c31676ff4672928b"), "Firstname" : "Sohan", "Middlename" : "Raj", "Lastname" : " M", "Age" : 30, "Salary" : 65000, "Designation" : "UI designer", "Role" : ["Team Lead", "Ui designer"] }
{ "_id" : ObjectId("62959166c31676ff4672928c"), "Firstname" : "Prateek", "Middlename" : "P", "Lastname" : "Nayak", "Age" : 24, "Salary" : 70000, "Designation" : "Manager", "Role" : ["Manager", "Tester", "Team Lead"] }
{ "_id" : ObjectId("6295924bc31676ff4672928d"), "Firstname" : "Sriman", "Middlename" : "S", "Lastname" : " Kumar", "Age" : 26, "Salary" : 75000, "Designation" : "UI designer", "Role" : ["Manager", "Ui designer"] }
{ "_id" : ObjectId("6295924bc31676ff4672928e"), "Firstname" : "Vishal", "Middlename" : "Kumar", "Lastname" : "H k", "Age" : 24, "Salary" : 72000, "Designation" : "Team Lead", "Role" : ["Manager", "Tester", "Team Lead"] }
{ "_id" : ObjectId("629592ecc31676ff4672928f"), "Firstname" : "Ullaas", "Middlename" : "M", "Lastname" : " Kumar", "Age" : 26, "Salary" : 20000, "Designation" : "UI designer", "Role" : ["Tester", "UI designer"] }
{ "_id" : ObjectId("629592ecc31676ff46729290"), "Firstname" : "Sanjeev", "Middlename" : "Sai", "Lastname" : "V V S", "Age" : 25, "Salary" : 32000, "Designation" : "Team Lead", "Role" : ["Manager", "Tester", "Team Lead"] }
{ "_id" : ObjectId("62959353c31676ff46729291"), "Firstname" : "Baba", "Middlename" : "M", "Lastname" : "Vinayak", "Age" : 35, "Salary" : 65000, "Designation" : "UI designer", "Role" : ["Tester", "UI designer"] }

{ "_id" : ObjectId("62959353c31676ff46729292"), "Firstname" : "Mohan", "Middlename" : "Kumar", "Lastname" : "T G", "Age" : 25, "Salary" : 30920, "Designation" : "Team Lead", "Role" : ["Manager", "Tester", "Team Lead"] }
{ "_id" : ObjectId("62959441c31676ff46729293"), "Firstname" : "Rajesh", "Middlename" : "Kumar", "Lastname" : "Rai", "Age" : 37, "Salary" : 75000, "Designation" : "Scientist", "Role" : ["Tester", "UI designer", "Manager"] }
{ "_id" : ObjectId("62959441c31676ff46729294"), "Firstname" : "Madhu", "Middlename" : "Rai", "Lastname" : "S G", "Age" : 25, "Salary" : 67920, "Designation" : "Team Lead", "Role" : ["Manager", "Team Lead"] }
{ "_id" : ObjectId("62959515c31676ff46729295"), "Firstname" : "Sachin", "Middlename" : "Ramesh", "Lastname" : "Tendulkar", "Age" : 45, "Salary" : 75000, "Designation" : "Scientist", "Role" : ["Tester", "UI designer", "Manager", "Team Lead"] }
{ "_id" : ObjectId("62959515c31676ff46729296"), "Firstname" : "Sunil", "Middlename" : "Kumar", "Lastname" : "Narine", "Age" : 30, "Salary" : 67920, "Designation" : "Team Lead", "Role" : ["Manager", "Team Lead", "Software developer"] }
{ "_id" : ObjectId("629595a5c31676ff46729297"), "Firstname" : "Laisha", "Middlename" : "sheik", "Lastname" : "abdul", "Age" : 23, "Salary" : 40000, "Designation" : "Tester", "Role" : ["Tester", "UI designer"] }
{ "_id" : ObjectId("629595a5c31676ff46729298"), "Firstname" : "Priya", "Middlename" : "Arul", "Lastname" : "Mohan", "Age" : 26, "Salary" : 67220, "Designation" : "Team Lead", "Role" : ["Manager", "Team Lead", "Software developer"] }



2)List all the records having salary in the range of 20000 – 35000

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

(or)

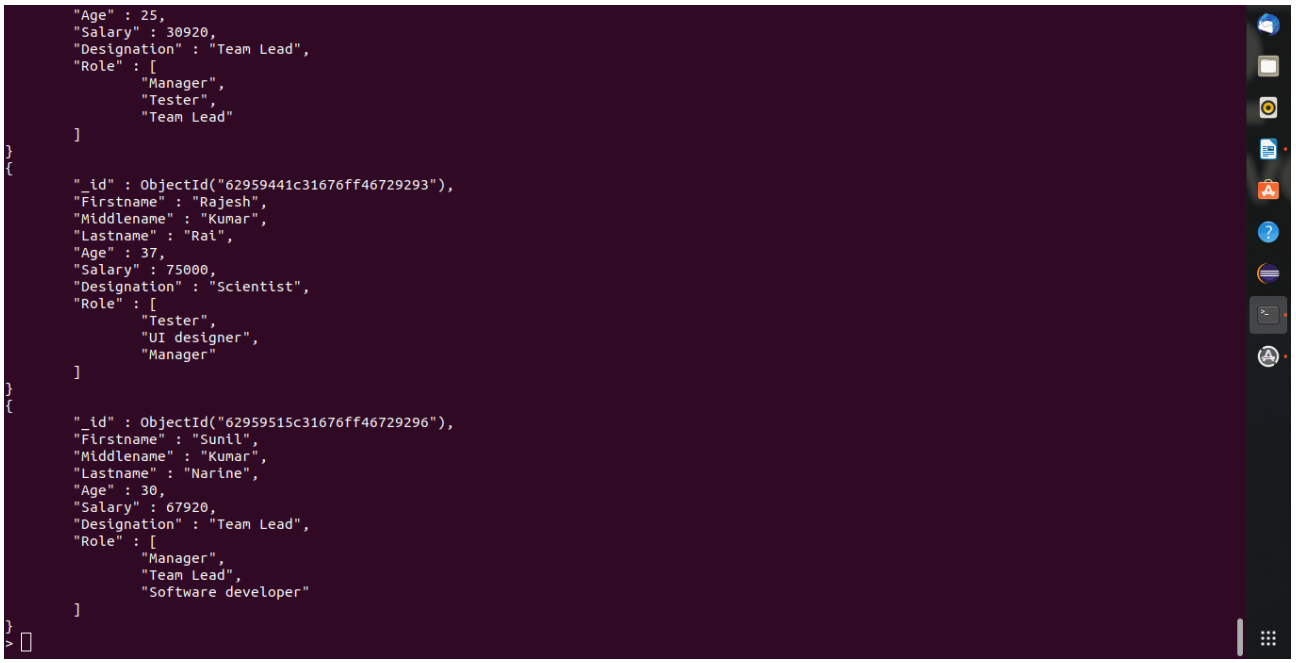
db.Employee.find({$and:[{Salary:{$gt:20000}},{Salary:{$lt:35000}}]})

```
>
> db.Employee.find({$and:[{Salary:{$gt:20000}},{Salary:{$lt:35000}}]})
{ "_id" : ObjectId("629592ecc31676ff46729290"), "Firstname" : "Sanjeev", "Middlename" : "Sai", "Lastname" : "V V S", "Age" : 25, "Salary" : 3200
0, "Designation" : "Team Lead", "Role" : [ "Manager", "Tester", "Team Lead" ] }
{ "_id" : ObjectId("62959353c31676ff46729292"), "Firstname" : "Mohan", "Middlename" : "Kumar", "Lastname" : "T G", "Age" : 25, "Salary" : 30920,
 "Designation" : "Team Lead", "Role" : [ "Manager", "Tester", "Team Lead" ] }
> db.Employee.find({$and:[{Salary:{$gt:20000}},{Salary:{$lt:35000}}]}).pretty()
{
        "_id" : ObjectId("629592ecc31676ff46729290"),
        "Firstname" : "Sanjeev",
        "Middlename" : "Sai",
        "Lastname" : "V V S",
        "Age" : 25,
        "Salary" : 32000,
        "Designation" : "Team Lead",
        "Role" : [
                "Manager",
                "Tester",
                "Team Lead"
        ]
}
{
        "_id" : ObjectId("62959353c31676ff46729292"),
        "Firstname" : "Mohan",
        "Middlename" : "Kumar",
        "Lastname" : "T G",
        "Age" : 25,
        "Salary" : 30920,
        "Designation" : "Team Lead",
        "Role" : [
                "Manager",
                "Tester",
                "Team Lead"
        ]
}
>
```

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

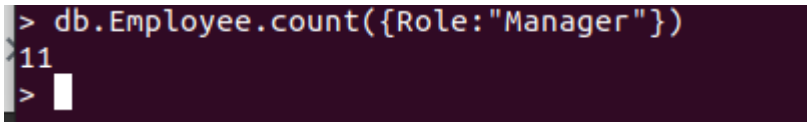
db.Employee.find({Middlename:"Kumar"})
(or)
db.Employee.find({Middlename:"Kumar"}).pretty()

```
> db.Employee.find({Middlename:"Kumar"})
{ "_id" : ObjectId("6295924bc31676ff4672928e"), "Firstname" : "Vishal", "Middlename" : "Kumar", "Lastname" : "H k", "Age" : 24, "Salary" : 72000
, "Designation" : "Team Lead", "Role" : [ "Manager", "Tester", "Team Lead" ] }
{ "_id" : ObjectId("62959353c31676ff46729292"), "Firstname" : "Mohan", "Middlename" : "Kumar", "Lastname" : "T G", "Age" : 25, "Salary" : 30920,
 "Designation" : "Team Lead", "Role" : [ "Manager", "Tester", "Team Lead" ] }
{ "_id" : ObjectId("62959441c31676ff46729293"), "Firstname" : "Rajesh", "Middlename" : "Kumar", "Lastname" : "Rai", "Age" : 37, "Salary" : 75000
, "Designation" : "Scientist", "Role" : [ "Tester", "UI designer", "Manager" ] }
{ "_id" : ObjectId("62959515c31676ff46729296"), "Firstname" : "Sunil", "Middlename" : "Kumar", "Lastname" : "Narine", "Age" : 30, "Salary" : 679
20, "Designation" : "Team Lead", "Role" : [ "Manager", "Team Lead", "Software developer" ] }
> db.Employee.find({Middlename:"Kumar"}).pretty()
{
        "_id" : ObjectId("6295924bc31676ff4672928e"),
        "Firstname" : "Vishal",
        "Middlename" : "Kumar",
        "Lastname" : "H k",
        "Age" : 24,
        "Salary" : 72000,
        "Designation" : "Team Lead",
        "Role" : [
                "Manager",
                "Tester",
                "Team Lead"
        ]
}
{
        "_id" : ObjectId("62959353c31676ff46729292"),
        "Firstname" : "Mohan",
        "Middlename" : "Kumar",
        "Lastname" : "T G",
        "Age" : 25,
        "Salary" : 30920,
        "Designation" : "Team Lead",
        "Role" : [
                "Manager",
                "Tester",
```



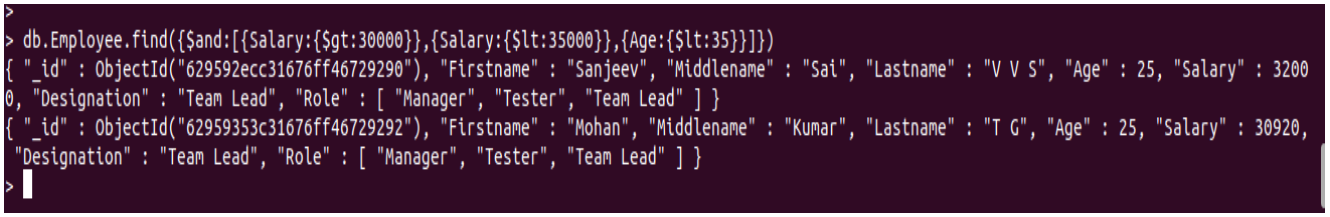
4)Count the number of Employees who has a role "Manager" in the Role field

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



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("629592ecc31676ff46729290"), "Firstname" : "Sanjeev", "Middlename" : "Sai", "Lastname" : "V V S", "Age" : 25, "Salary" : 3200
0, "Designation" : "Team Lead", "Role" : [ "Manager", "Tester", "Team Lead" ] }
{ "_id" : ObjectId("62959353c31676ff46729292"), "Firstname" : "Mohan", "Middlename" : "Kumar", "Lastname" : "T G", "Age" : 25, "Salary" : 30920,
 "Designation" : "Team Lead", "Role" : [ "Manager", "Tester", "Team Lead" ] }
>
```

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

db.Employee.remove({$and:[{Firstname:"Rajesh"},{Designation:"Scientist"}]})

```
> db.Employee.find({Firstname:"Rajesh"})
{ "_id" : ObjectId("62959089c31676ff4672928a"), "Firstname" : "Rajesh", "Middlename" : "C", "Lastname" : "Rajesh", "Age" : 26, "Salary" : 50000,
 "Designation" : "Tester", "Role" : [ "Manager", "Tester", "UI designer" ] }
{ "_id" : ObjectId("62959441c31676ff46729293"), "Firstname" : "Rajesh", "Middlename" : "Kumar", "Lastname" : "Rai", "Age" : 37, "Salary" : 75000
, "Designation" : "Scientist", "Role" : [ "Tester", "UI designer", "Manager" ] }
> db.Employee.remove({$and:[{Firstname:"Rajesh"},{Designation:"Scientist"}]})
WriteResult({ "nRemoved" : 1 })
> db.Employee.find({Firstname:"Rajesh"})
{ "_id" : ObjectId("62959089c31676ff4672928a"), "Firstname" : "Rajesh", "Middlename" : "C", "Lastname" : "Rajesh", "Age" : 26, "Salary" : 50000,
 "Designation" : "Tester", "Role" : [ "Manager", "Tester", "UI designer" ] }
>
```

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({Designation:"Team Lead"})
{ "_id" : ObjectId("6295924bc31676ff4672928e"), "Firstname" : "Vishal", "Middlename" : "Kumar", "Lastname" : "H k", "Age" : 24, "Salary" : 72000
, "Designation" : "Team Lead", "Role" : [ "Manager", "Tester", "Team Lead" ] }
{ "_id" : ObjectId("629592ecc31676ff46729290"), "Firstname" : "Sanjeev", "Middlename" : "Sai", "Lastname" : "V V S", "Age" : 25, "Salary" : 3200
0, "Designation" : "Team Lead", "Role" : [ "Manager", "Tester", "Team Lead" ] }
{ "_id" : ObjectId("62959353c31676ff46729292"), "Firstname" : "Mohan", "Middlename" : "Kumar", "Lastname" : "T G", "Age" : 25, "Salary" : 30920,
 "Designation" : "Team Lead", "Role" : [ "Manager", "Tester", "Team Lead" ] }
{ "_id" : ObjectId("62959441c31676ff46729294"), "Firstname" : "Madhu", "Middlename" : "Rai", "Lastname" : "S G", "Age" : 25, "Salary" : 67920, "
Designation" : "Team Lead", "Role" : [ "Manager", "Team Lead" ] }
{ "_id" : ObjectId("62959515c31676ff46729296"), "Firstname" : "Sunil", "Middlename" : "Kumar", "Lastname" : "Narine", "Age" : 30, "Salary" : 679
20, "Designation" : "Team Lead", "Role" : [ "Manager", "Team Lead", "Software developer" ] }
{ "_id" : ObjectId("629595a5c31676ff46729298"), "Firstname" : "Priya", "Middlename" : "Arul", "Lastname" : "Mohan", "Age" : 26, "Salary" : 67220
, "Designation" : "Team Lead", "Role" : [ "Manager", "Team Lead", "Software developer" ] }
> db.Employee.updateMany({Role:"Team Lead"},{$set:{Salary:55650}})
{ "acknowledged" : true, "matchedCount" : 10, "modifiedCount" : 10 }
> db.Employee.find({Designation:"Team Lead"})
{ "_id" : ObjectId("6295924bc31676ff4672928e"), "Firstname" : "Vishal", "Middlename" : "Kumar", "Lastname" : "H k", "Age" : 24, "Salary" : 55650
, "Designation" : "Team Lead", "Role" : [ "Manager", "Tester", "Team Lead" ] }
{ "_id" : ObjectId("629592ecc31676ff46729290"), "Firstname" : "Sanjeev", "Middlename" : "Sai", "Lastname" : "V V S", "Age" : 25, "Salary" : 5565
0, "Designation" : "Team Lead", "Role" : [ "Manager", "Tester", "Team Lead" ] }
{ "_id" : ObjectId("62959353c31676ff46729292"), "Firstname" : "Mohan", "Middlename" : "Kumar", "Lastname" : "T G", "Age" : 25, "Salary" : 55650,
 "Designation" : "Team Lead", "Role" : [ "Manager", "Tester", "Team Lead" ] }
{ "_id" : ObjectId("62959441c31676ff46729294"), "Firstname" : "Madhu", "Middlename" : "Rai", "Lastname" : "S G", "Age" : 25, "Salary" : 55650, "
Designation" : "Team Lead", "Role" : [ "Manager", "Team Lead" ] }
{ "_id" : ObjectId("62959515c31676ff46729296"), "Firstname" : "Sunil", "Middlename" : "Kumar", "Lastname" : "Narine", "Age" : 30, "Salary" : 556
50, "Designation" : "Team Lead", "Role" : [ "Manager", "Team Lead", "Software developer" ] }
{ "_id" : ObjectId("629595a5c31676ff46729298"), "Firstname" : "Priya", "Middlename" : "Arul", "Lastname" : "Mohan", "Age" : 26, "Salary" : 55650
, "Designation" : "Team Lead", "Role" : [ "Manager", "Team Lead", "Software developer" ] }
> -
```

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" : 30, "Average" : 55650 }
{ "_id" : 26, "Average" : 50162.5 }
{ "_id" : 24, "Average" : 55650 }
{ "_id" : 35, "Average" : 65000 }
{ "_id" : 45, "Average" : 55650 }
{ "_id" : 25, "Average" : 55650 }
{ "_id" : 23, "Average" : 47825 }
>
```

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()