VISVESVARAYA TECHNOLOGICAL UNIVERSITY

"JnanaSangama", Belgaum -590014, Karnataka.



LAB REPORT on

BIG DATA ANALYTICS (20CS6PEBDA)

Submitted by

KIRAN M K (1BM19CS073)

in partial fulfillment for the award of the degree of BACHELOR OF ENGINEERING
in
COMPUTER SCIENCE AND ENGINEERING



B.M.S. COLLEGE OF ENGINEERING
(Autonomous Institution under VTU)
BENGALURU-560019
May-2022 to July-2022

B. M. S. College of Engineering,

Bull Temple Road, Bangalore 560019
(Affiliated To Visvesvaraya Technological University, Belgaum)

Department of Computer Science and Engineering



CERTIFICATE

This is to certify that the Lab work entitled "BIG DATA ANALYTICS" carried out by KIRAN M K(1BM19CS073), who is bonafide student of B. M. S. College of Engineering. It is in partial fulfillment for the award of Bachelor of Engineering in Computer Science and Engineering of the Visvesvaraya Technological University, Belgaum during the year 2022. The Lab report has been approved as it satisfies the academic requirements in respect of a Big Data Analytics Laboratory - (20CS6PEBDA) work prescribed for the said degree.

Dr. Pallavi G. B.Assistant Professor
Department of CSE
BMSCE, Bengaluru

Dr. Jyothi S Nayak Professor and Head Department of CSE BMSCE, Bengaluru

Index Sheet

Sl. No.	Experiment Title	Page No.
1	<u>Cassandra Lab Program 1: -</u> Create a Data set either structured/Semi-Structured/Unstructured from Twitter/Facebook etc. to perform various DB operations using Cassandra. (Use the Face Pager app to perform real-time streaming)	1-3
2	<u>Cassandra Lab Program 2: -</u> Create a Data set either structured/Semi-Structured/Unstructured from Twitter/Facebook etc. to perform various DB operations using Cassandra. (Use the Face Pager app to perform real-time streaming)	4-6
3	MongoDB Lab Program 1 (CRUD Demonstration): - Students should be classifying a dataset into one of the standard forms and apply suitable querying rules to obtain suitable results	7-11
4	MongoDB Lab Program 2 (CRUD Demonstration): - Students should be classifying a dataset into one of the standard forms and apply suitable querying rules to obtain suitable results	12-17

Course Outcome

CO1	Apply the concept of NoSQL, Hadoop or Spark for a given task
CO2	Analyze the Big Data and obtain insight using data analytics mechanisms.
CO3	Design and implement Big data applications by applying NoSQL, Hadoop or Spark

Cassandra Program 1

Program 1. Perform the following DB operations using Cassandra.

1. Create a key space by name Employee

```
cqlsh> CREATE KEYSPACE employee WITH replication = {'class':'SimpleStrategy', 'replication_factor':1};
cqlsh> describe keyspace
No keyspace specified and no current keyspace
cqlsh> describe Employee;

CREATE KEYSPACE employee WITH replication = {'class': 'SimpleStrategy', 'replication_factor': '1'} AND durable_writes = true;
cqlsh> create table Employee.employee_info(emp_id int Primary Key, emp_name text, designation text, date_of_joining timestamp, salary double,
dept_name text);
cqlsh> select * from Employee.employee_info;
emp_id | date_of_joining | dept_name | designation | emp_name | salary
```

2. Create a column family by name Employee-Info with attributes Emp_Id Primary Key, Emp_Name,

Designation, Date_of_Joining, Salary, Dept_Name

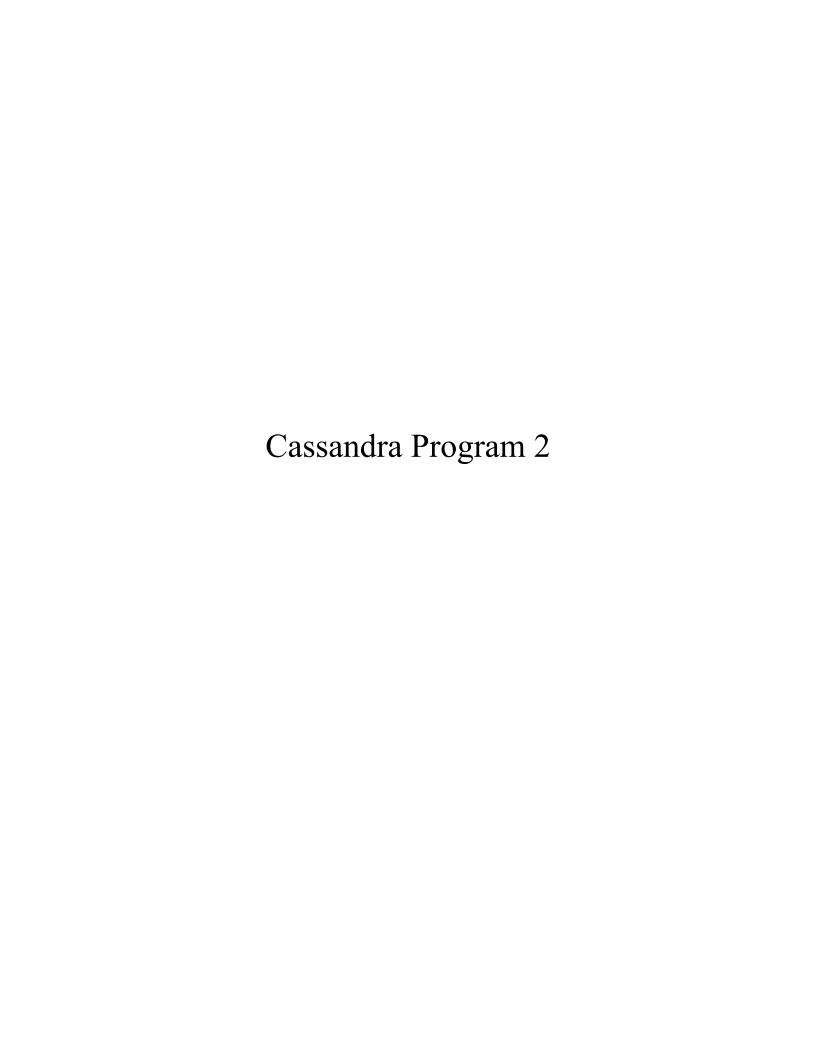
3. Insert the values into the table in batch

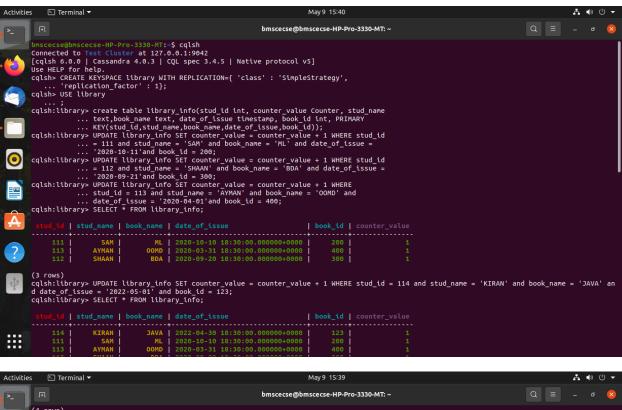
4. Update Employee name and Department of Emp-Id 121

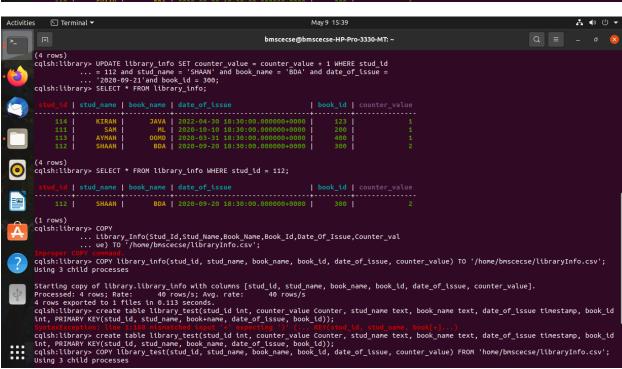
5. Sort the details of Employee records based on salary

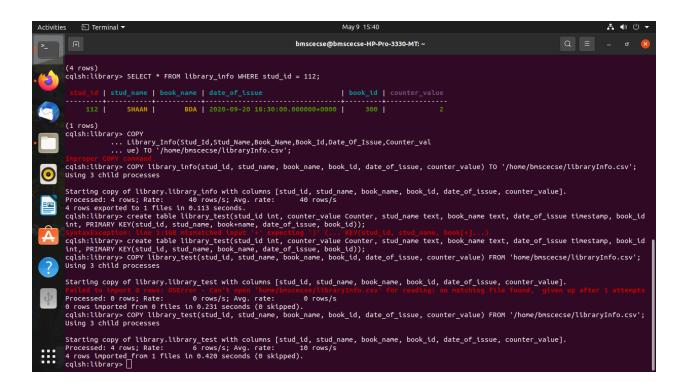
6. Alter the schema of the table Employee_Info to add a column Projects which stores a set of Projects done by the corresponding Employee.

7. Update the altered table to add project names.









Mongo DB Program 1

```
| Description |
```

```
db.Student.find({StudName:/M/}).pretty();
   "id" : 1, "StudName" : "Kiran M K", "sem" : 6, "Hobbies" : "Podcasts" }
db.Student.find({StudName:/M/i}).pretty();
   "_id" : 1, "StudName" : "Kiran M K", "sem" : 6, "Hobbies" : "Podcasts" }

   "_id" : 9,
    "StudName" : "Trivikram hegde",
   "sem" : 6,
   "Hobbies" : "Chess"

   "_id" : 10,
   "StudName" : "Padmanabha Shetty",
   "sem" : 6,
   "Hobbies" : "Skating"

db.Student.find({StudName:/[mM]/}).pretty();
   "_id" : 1, "StudName" : "Kiran M K", "sem" : 6, "Hobbies" : "Podcasts" }

   "_id" : 9,
   "StudName" : "Trivikram hegde",
   "sem" : 6,
   "Hobbies" : "Chess"

   "_id" : 10,
   "StudName" : "Padmanabha Shetty",
   "sem" : 6,
   "Hobbies" : "Skating"

db.Student.find({Hobbies:{$eq:'Chess'}}).pretty();
```

```
db.Student.find({_id:1},{StudName:1,sem:1,_id:0});
"StudName" : "Kiran", "sem" : 6 }
 db.Student.find({Hobbies:{$ne:'Chess'}}).pretty();
  "_id" : 1, "StudName" : "Kiran", "sem" : 6, "Hobbies" : "Podcasts" }
        "_id" : 2,
"StudName" : "Trivikram Hegde",
         "sem" : 6,
         "Hobbies" : "Political analysis"
        "_id" : 4,
"StudName" : "Laran",
         "sem" : 6,
"Hobbies" : "Skating",
         "Location" : "Bengaluru"
 "_id" : ObjectId("62543111e29a7583f87aa599"), "StudName" : "Kiran" }
> db.Student.update({_id:3},{$set:{Location:null}});
WriteResult({ "nMatched" : 1, "nUpserted" : 0, "nModified" : 1 })
> db.Student.count();
> db.Student.find({sem:6}).limit(3).pretty();
 "_id" : 1, "StudName" : "Kiran", "sem" : 6, "Hobbies" : "Podcasts" }
         "_id" : 2,
"StudName" : "Trivikram Hegde",
         "sem" : 6,
"Hobbies" : "Political analysis"
         "_id" : 3,
         "StudName": "Padmanabha Shetty",
         "sem" : 6,
         "Hobbies" : "Chess",
"Location" : null
```

```
> db.Student.count();
4
    db.Student.find().sort({StudName:-1}).pretty();
{
        "_id" : 2,
        "StudName" : "Trivikram Hegde",
        "sem" : 6,
        "Hobbies" : "Political analysis"
}
{
        "_id" : 3,
        "StudName" : "Padmanabha Shetty",
        "sem" : 6,
        "Hobbies" : "Chess"
}
{
        "_id" : 4, "StudName" : "Laran", "sem" : 6, "Hobbies" : "Skating" }
{
        "_id" : 4, "StudName" : "Kiran", "sem" : 6, "Hobbies" : "Podcasts" }
> db.Student.save({StudName:'Kiran"})
WriteResult({ "Inserted" : 1 })
> db.Student.find();
{
        "_id" : 2, "StudName" : "Kiran", "sem" : 6, "Hobbies" : "Podcasts" }
{
        "_id" : 3, "StudName" : "Kiran", "sem" : 6, "Hobbies" : "Political analysis" }
{
        "_id" : 3, "StudName" : "Rivikram Hegde", "sem" : 6, "Hobbies" : "Chess" }
{
        "_id" : 4, "StudName" : "Laran", "sem" : 6, "Hobbies" : "Skating" }
{
        "_id" : 0bjectId("62543111e29a7583187aa599"), "StudName" : "Kiran" }
> db.Student.update([ td:4), (Sset:(Location:"Bengaluru")));
WriteResult({ "nMatched" : 1, "nUpserted" : 0, "nModified" : 1 })
> db.Student.find();
{
        "_id" : 1, "StudName" : "Kiran", "sem" : 6, "Hobbies" : "Podcasts" }
{
        "_id" : 2, "StudName" : "Kiran", "sem" : 6, "Hobbies" : "Political analysis" }
{
        "_id" : 3, "StudName" : "Frivikram Hegde", "sem" : 6, "Hobbies" : "Political analysis" }
{
        "_id" : 3, "StudName" : "Frivikram Hegde", "sem" : 6, "Hobbies" : "Chess" }
{
        "_id" : 4, "StudName" : "Frivikram Hegde", "sem" : 6, "Hobbies" : "Chess" }
{
        "_id" : 3, "StudName" : "Padmanabha Shetty", "sem" : 6, "Hobbies" : "Chess" }
{
        "_id" : 4, "StudName" : "Rivan", "sem" : 6, "Hobbies" : "Skating", "Location" : "Bengaluru" }
{
        "_id" : 4, "StudName" : "Rivan", "sem" : 6, "Hobbies" : "Kiran" }
}
{
        "_id" : 4, "StudName" : "Rivan", "sem" : 6, "Hobbies" : "Skating", "Location" : "Bengaluru" }
{
        "_id" : 4, "StudName" : "Rivan", "sem" : 6, "Hobbies" : "Kiran" }
}
{
        "_id" : 4, "StudName" :
```

```
db.Student.find().sort({StudName:1}).pretty();
"_id" : 1, "StudName" : "Kiran", "sem" : 6, "Hobbies" : "Podcasts" }
"_id" : ObjectId("62543111e29a7583f87aa599"), "StudName" : "Kiran" }
         "_id" : 4,
"StudName" : "Laran",
         "sem" : 6,
"Hobbies" : "Skating",
"Location" : "Bengaluru"
         "_id" : 3,
"StudName" : "Padmanabha Shetty",
          "sem" : 6,
         "Hobbies" : "Chess",
         "Location" : null
         "_id" : 2,
"StudName" : "Trivikram Hegde",
          "sem" : 6,
"Hobbies" : "Political analysis"
db.Student.find().skip(2).pretty();
         "_id" : 3,
"StudName" : "Padmanabha Shetty",
          "sem" : 6,
         "Hobbies": "Chess",
         "Location" : null
         "_id" : 4,
"StudName" : "Laran",
         "sem" : 6,
"Hobbies" : "Skating",
"Location" : "Bengaluru"
```

```
db.food.insertOne( { _id: 5, fruits: [ "mango", "apple", "orange" ] } );
"acknowledged" : true, "insertedId" : 5 }
db.food.updateOne( { _id: 5 }, { $pop: { fruits: -1 } } );
"acknowledged" : true, "matchedCount" : 1, "modifiedCount" : 1 }
    db.food.find();
  "_id" : 1, "fruits" : [ "grapes", "mango", "apple" ] }
"_id" : 2, "fruits" : [ "grapes", "mango", "cherry" ], "price" : [ { "grapes" : 80, "mango" : 200, "cherry" : 100 } ] }
"_id" : 3, "fruits" : [ "banana", "apple" ] }
"_id" : 5, "fruits" : [ "apple", "orange" ] }
   db.createCollection("Customers"):
"_id" : 2, "TotAccBal" : 38000 }
    "_id" : 1, "TotAccBal" : 31000 }
    db.Customers.insert({custID:s, AccBal:10000000, AccType: "Current"});
IriteResult({ "nInserted" : 1 })
    db.Customers.aggregate( {Smatch:{AccType:"Savings"}},{$group : { _id : "$custID",TotAccBal : {$sum:"$AccBal"} } });
    db.Customers.aggregate( {Smatch:{AccType:"Savings"}},{$group : { _id : "$custID",TotAccBal : {$sum:"$AccBal"} } });
    db.Customers.aggregate( {Smatch:{AccType:"Savings"}},{$group : { _id : "$custID",TotAccBal : {$sum:"$AccBal"} } });
    "_id" : 2, "TotAccBal" : 38000 }
    "_id" : 1, "TotAccBal" : 31000 }
    db.Customers.find();
    "_id" : 1, "custID" : 1, "AccBal" : 15000, "AccType" : "Savings" }
    "_id" : 2, "custID" : 1, "AccBal" : 16000, "AccType" : "Savings" }
    "_id" : 2, "custID" : 2, "AccBal" : 20000, "AccType" : "Savings" }
    "_id" : 0bjectId("625448f0ee5e75021943909a"), "custID" : 5, "AccBal" : 10000000, "AccType" : "Current" }
    db.Customers.find();
    "_id" : 1, "custID" : 1, "AccBal" : 15000, "AccType" : "Savings" }
    "_id" : 1, "custID" : 1, "AccBal" : 15000, "AccType" : "Savings" }
    "_id" : 1, "custID" : 1, "AccBal" : 16000, "AccType" : "Savings" }
    "_id" : 2, "custID" : 1, "AccBal" : 16000, "AccType" : "Savings" }
    "_id" : 3, "custID" : 2, "AccBal" : 18000, "AccType" : "Savings" }
    "_id" : 3, "custID" : 2, "AccBal" : 18000, "AccType" : "Savings" }
    "_id" : 4, "custID" : 2, "AccBal" : 18000, "AccType" : "Savings" }
    "_id" : 4, "custID" : 2, "AccBal" : 18000, "AccType" : "Savings" }
    "_id" : 0bjectId("625448f0ee5e75021943909a"), "custID" : 5, "AccBal" : 10000000, "AccType" : "Current" }
    "_id" : 0bjectId("62544974ee5e75021943909a"), "custID" : 5, "AccBal" : 10000000, "AccType" : "Savings" }
    "_id" : 0bjectId("62544974ee5e75021943909a"), "custID" : 3, "AccBal" : 10000000, "AccType" : "Savings" }
    "_id" : 0bjectId("62544974ee5e75021943909a"), "custID" : 3, "AccBal" : 10000000, "AccType" : "Savings" }
    "_id" : 0bjectId("62544974ee5e75021943909a"), "cus
     db.food.updateOne({_id:1},{$addToSet:{fruits:"banana"}});
"acknowledged" : true, "matchedCount" : 1, "modifiedCount" : 1 }
     db.food.find();
    db.food.find();
   "id" : 1, "fruits" : [ "grapes", "mango", "apple", "banana" ] }
   "_id" : 2, "fruits" : [ "grapes", "mango", "cherry" ], "price" : [ { "grapes" : 80, "mango" : 200, "cherry" : 100 } ] }
   "id" : 3, "fruits" : [ "banana", "apple" ] }
   "_id" : 5, "fruits" : [ "apple", "orange" ] }
   db.food.updateOne({_id:1}, {$pullAll:{fruits:["mango", "apple"]}});
   "acknowledged" : true, "matchedCount" : 1, "modifiedCount" : 1 }
}
 "acknowledged": true, "Matchedcount".,
db.food.find();

"_id": 1, "fruits": [ "grapes", "banana" ] }

"_id": 2, "fruits": [ "grapes", "mango", "cherry" ], "price": [ { "grapes": 80, "mango": 200, "cherry": 100 } ] }

"_id": 3, "fruits": [ "banana", "apple" ] }

"_id": 5, "fruits": [ "apple", "orange" ] }
db.food.updateMany({},{$pull:{fruits:{$in:["mango", "grapes"]}}});

"acknowledged": true, "matchedCount": 4, "modifiedCount": 2 }
```

db.food.find();
 "_id" : 1, "fruits" : ["banana"] }
 "_id" : 2, "fruits" : ["cherry"], "price" : [{ "grapes" : 80, "mango" : 200, "cherry" : 100 }] }
 "_id" : 3, "fruits" : ["banana", "apple"] }
 "_id" : 5, "fruits" : ["apple", "orange"] }

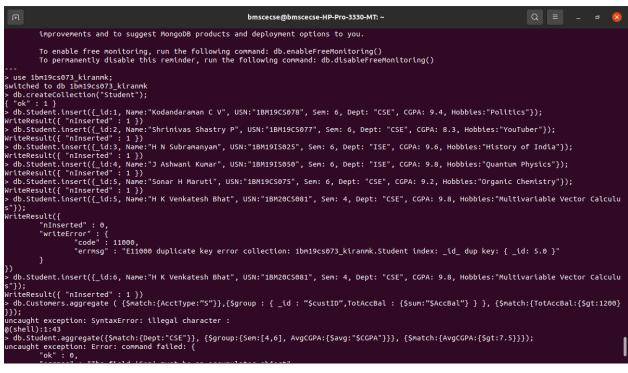


```
bmscecse@bmscecse-HP-Pro-3330-MT: ~
 bmscecse@bmscecse-HP-Pro-3330-MT:-$ sudo mongoexport --host localhost --db 1bm19cs073_kiranmk --collection Student --type=csv --out /home/bmsce/Desktop/output.txt --fields "_id","Name","USN","Sem","Dept","CGPA","Hobbies"
2022-04-18T15:13:05.991+0530 connected to: mongodb://localhost/
2022-04-18T15:13:05.996+0530 exported 6 records
2022-04-18T15:13:05.991+0530
2022-04-18T15:13:05.996+0530
2022-04-18T15:13:05.996-0530 exported 6 records
bmscecsegbmscecse-HP-Pro-3330-MT:-5 ls

024 Desktop Documents Downloads Music Pictures Public Templates Videos
bmscecsegbmscecse-HP-Pro-3330-MT:-5 cd Desktop
bmscecsegbmscecse-HP-Pro-3330-MT:-/Desktop$ ls
bmscecsegbmscecse-HP-Pro-3330-MT:-/Desktop$ mongoexport --host localhost --db 1bm19cs073_kiranmk --collection Student --csv --out /home/bmsce
/Desktop/output.txt --fields "_id", "Name", "USN", "Sem", "Dept", "CGPA", "Hobbies"

2022-04-18T15:14:16.325+0530 connected to: mongodb://localhost/
2022-04-18T15:14:16.325+0530 error opening output stream: mkdir /home/bmsce/Desktop: permission denied
bmscecsegbmscecse-HP-Pro-3330-MT:-/Desktop$ sudo mongoexport --host localhost --db 1bm19cs073_kiranmk --collection Student --csv --out /home/
bmsce/Desktop/output.txt --fields "_id", "Name", "USN", "Sem", "Dept", "CCPA", "Hobbies"

2022-04-18T15:14:22.162+0530 csv flag is deprecated; please use --type=csv instead
connected to: mongodb://localhost/
2022-04-18T15:14:22.164+0530 csv flag is deprecated; please use --type=csv instead
2022-04-18T15:14:22.164+0530 exported 6 records
2022-04-18T15:14:22.169+0530
                                                                          0530 exported 6 records
-3330-MT:~/Desktop$ pwd
  home/bmscecse/Desktop
/nome/pmscecse/Desktop
bmscecse/Desktop/sudo mongoexport --host localhost --db 1bm19cs073_kiranmk --collection Student --type=csv --out /
home/bmscecse/Desktop/output.txt --fields "_id","Name","USN","Sem","Dept","CGPA","Hobbies"
2022-04-18T15:15:10.012-0530
2022-04-18T15:15:10.017-0530
exported 6 records
bmscecse@bmscecse-HP-Pro-3330-NT:-/Desktop$ cd ..
bmscecse@bmscecse-HP-Pro-3330-NT:-/Desktop$ cd ..
    nscecse@bmscecse-HP-Pro-3330-MT:~/Desk
nscecse@bmscecse-HP-Pro-3330-MT:~$ ls
    24 Desktop Documents Downloads Music
mscecse@bmscecse-HP-Pro-3330-MT:/\square\text{scd}...
mscecse@bmscecse-HP-Pro-3330-MT:/\home$ ls
                                  scecse-HP-Pro-3330-MT:/home$ cd bmsce
   ash: cd: bmsce: Permission denied
                                                                                     0-MT:/home$ sudo cd bmsce
 sudo: cd: command not found
                                                                             3330-MT:/home$ login bmsce
login: Cannot possibly work without effective root baseecse@bmscecse-HP-Pro-3330-MT:/home$ cd bmscecsebmscecse@bmscecse-HP-Pro-3330-MT:-$
```



```
Improvements and to suggest MongoDB products and deployment options to you.

To enable free monitoring, run the following command: db.enableFreeMonitoring()
To permanently disable this reminder, run the following command: db.disableFreeMonitoring()

> use ibmigcs073_ktrannk;
switched to db ibmigcs073_ktrannk

> db.creatcollection("Student");
{ "ok" : 1 }

> db.student.insert([id:1, Name: "Kodandaraman C V", USN: "IBMI9CS078", Sem: 6, Dept: "CSE", CGPA: 9.4, Hobbies: "Politics"));
WriteResult(("inInserted": 1 ))

> db.student.insert([id:2, Name: "Shrinivas Shastry P", USN: "IBMI9CS077", Sem: 6, Dept: "CSE", CGPA: 9.4, Hobbies: "YouTuber"));
WriteResult(("inInserted": 1 ))

> db.student.insert([id:4, Name: "J Ashwani Kumar", USN: "IBMI9CS077", Sem: 6, Dept: "ISE", CGPA: 9.6, Hobbies: "Multivariable Vector Calculu Student insert([id:4, Name: "J Ashwani Kumar", USN: "IBMI9CS075", Sem: 6, Dept: "CSE", CGPA: 9.8, Hobbies: "Quantum Physics"));
WriteResult(("inInserted": 1 ))

> db.student.insert([id:5, Name: "K Venkatesh Bhat", USN: "IBMI9CS075", Sem: 4, Dept: "CSE", CGPA: 9.8, Hobbies: "Multivariable Vector Calculu Student insert([id:5, Name: "K Venkatesh Bhat", USN: "IBM20CS081", Sem: 4, Dept: "CSE", CGPA: 9.8, Hobbies: "Multivariable Vector Calculu "Code": 11000,

"writeBroor": {
"Code": 11000,
"errnsg": "E11000 duplicate key error collection: 1bm19cs073_ktrannk.Student index: _td_ dup key: { _id: 5.0 }"

} ob.student.insert([id:6, Name: "H K Venkatesh Bhat", USN: "1BM20CS081", Sem: 4, Dept: "CSE", CGPA: 9.8, Hobbies: "Multivariable Vector Calculu S"));

WriteResult(("InInserted": 1))

> db.student.insert([id:6, Name: "H K Venkatesh Bhat", USN: "1BM20CS081", Sem: 4, Dept: "CSE", CGPA: 9.8, Hobbies: "Multivariable Vector Calculu S"));

WriteResult(("InInserted": 1))

> db.student.insert([id:6, Name: "H K Venkatesh Bhat", USN: "1BM20CS081", Sem: 4, Dept: "CSE", CGPA: 9.8, Hobbies: "Multivariable Vector Calculu S"));

WriteResult(("InInserted": 1))

> db.student.insert([id:6, Name: "H K Venkatesh Bhat", USN: "1BM20CS0
```

```
| RirangAlran-H310M-H-2-0:~ | Q | E - 0 | Name | Content | Content
```

