Name: Sumit Singh

Roll No: 380

Class: TYBSC CS A

Subject: Data Science

Practical No : 1
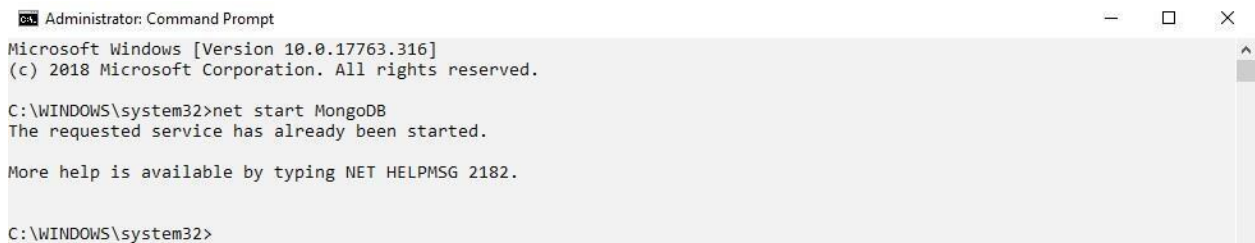
Aim: Practical of Data Collection, Data Curation and Management for large scale data system (such as MongoDB).

Procedure:

Install MongoDB

Run CMD as administrator

Start MongoDB via cmd

Enter the MongoDB bin folder via the cmd.

Type Mongo

```
Administrator: Command Prompt - mongo                                    —   □   ×

C:\WINDOWS\system32>cd C:\Program Files\MongoDB\Server\4.0\bin

C:\Program Files\MongoDB\Server\4.0\bin>mongo
MongoDB shell version v4.0.6
connecting to: mongodb://127.0.0.1:27017/?gssapiServiceName=mongodb
Implicit session: session { "id" : UUID("fef6c64f-6507-41d6-bed2-de47c928ec45") }
MongoDB server version: 4.0.6
Welcome to the MongoDB shell.
For interactive help, type "help".
For more comprehensive documentation, see
        http://docs.mongodb.org/
Questions? Try the support group
        http://groups.google.com/group/mongodb-user
Server has startup warnings:
2019-03-08T10:28:17.164+0530 I CONTROL  [initandlisten]
2019-03-08T10:28:17.165+0530 I CONTROL  [initandlisten] ** WARNING: Access control is not enabled for the database.
2019-03-08T10:28:17.165+0530 I CONTROL  [initandlisten] **          Read and write access to data and configuration is u
nrestricted.
2019-03-08T10:28:17.165+0530 I CONTROL  [initandlisten]
---
Enable MongoDB's free cloud-based monitoring service, which will then receive and display
metrics about your deployment (disk utilization, CPU, operation statistics, etc).

The monitoring data will be available on a MongoDB website with a unique URL accessible to you
and anyone you share the URL with. MongoDB may use this information to make product
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()
---
```

Check Current Database

```
> db
test
>
```

Use database library

```
> use library
switched to db library
>
```

Create Collection(table)

```
> db.createCollection("book_info")
{ "ok" : 1 }
>
```

Insert document (records) in collection

```
> db.book_info.insert({title:'Programming with Java',status_info:{accession_no:BS0001,status:ISSUES},)
...
...
> db.book_info.insert({title:'PROGRAMMING WITH JAVA',
... status_info:{accession_no:'BS0001',status:'ISSUED'},
... author:'EBALAGURUSWAMY',
... cost:'350',
... publisher:{name:'TMH',location:'banglore'}})
WriteResult({ "nInserted" : 1 })
```

To insert multiple records in one column

```
> db.book_info.insert({title:'DISTRIBUTED SYSTEMS', status_info:[{accession_no:'BS0005',status:'ISSUED'},{accession_no:
BS0006',status:'ISSUED'},{accesion_no:'BS0007',status:'ISSUED'},{accession_no:'BS0008',status:'AVAIL'}], author:'ANDREW
TANENBAUM', cost:'350', publisher:{name:'PEARSON',location:'ANDHERI'}})
WriteResult({ "nInserted" : 1 })
```

To view records of a specific document:

```
> db.book_info.find({title:'LET US C'})
{ "_id" : ObjectId("5c8201d69cff5f2a54951b57"), "title" : "LET US C", "status_info" : { "accession_no" : "BS0009", "stat
us" : "AVAIL" }, "author" : "KANETKAR YASHWANT P", "cost" : "600", "publisher" : { "name" : "B.P.B", "location" : "RAJAS
THAN" } }
>
```

To view all records:

```
> db.book_info.find({})
{ "_id" : ObjectId("5c81fd569cff5f2a54951b52"), "title" : "PROGRAMMING WITH JAVA", "status_info" : { "accession_no" : "B
S0001", "status" : "ISSUED" }, "author" : "EBALAGURUSWAMY", "cost" : "350", "publisher" : { "name" : "TMH", "location" :
 "banglore" } }
{ "_id" : ObjectId("5c81fdd89cff5f2a54951b53"), "title" : "ASP.NET3.5 VB 2008", "status_info" : { "accession_no" : "BS00
02", "status" : "AVAIL" }, "author" : "ANNE BOEHM", "cost" : "650", "publisher" : { "name" : "MURACH", "location" : "JAI
PUR" } }
{ "_id" : ObjectId("5c81fe5e9cff5f2a54951b54"), "title" : "PRPGRAMMING IN VB", "status_info" : { "accession_no" : "BS000
3", "status" : "ISSUED" }, "author" : "JULIA CASE BRADLEY", "cost" : "600", "publisher" : { "name" : "TMH", "location" :
 "BANGLORE" } }
{ "_id" : ObjectId("5c81feaf9cff5f2a54951b55"), "title" : "DATABASE SYSTEM CONCEPTS", "status_info" : { "accession_no" :
 "BS0004", "status" : "ISSUED" }, "author" : "KORTH SUDARSHAN", "cost" : "500", "publisher" : { "name" : "TMH", "locatio
n" : "BANGLORE" } }
{ "_id" : ObjectId("5c82005f9cff5f2a54951b56"), "title" : "DISTRIBUTED SYSTEMS", "status_info" : [ { "accession_no" : "B
S0005", "status" : "ISSUED" }, { "accession_no" : "BS0006", "status" : "ISSUED" }, { "accesion_no" : "BS0007", "status"
: "ISSUED" }, { "accession_no" : "BS0008", "status" : "AVAIL" } ], "author" : "ANDREW TANENBAUM", "cost" : "350", "publi
sher" : { "name" : "PEARSON", "location" : "ANDHERI" } }
{ "_id" : ObjectId("5c8201d69cff5f2a54951b57"), "title" : "LET US C", "status_info" : { "accession_no" : "BS0009", "stat
us" : "AVAIL" }, "author" : "KANETKAR YASHWANT P", "cost" : "600", "publisher" : { "name" : "B.P.B", "location" : "RAJAS
THAN" } }
{ "_id" : ObjectId("5c8202f69cff5f2a54951b58"), "title" : "MODERN DIGITAL ELECTRONICS", "status_info" : [ { "accession_n
o" : "BS010", "status" : "ISSUED" }, { "accession_no" : "BS011", "status" : "AVAIL" } ], "author" : "JAIN R.P", "cost" :
 "650", "publisher" : { "name" : "TMH", "location" : "BANGLORE" } }
```

To view accession numbers:

```
> db.book_info.find({},{'status_info.accession_no':1})
{ "_id" : ObjectId("5c81fd569cff5f2a54951b52"), "status_info" : { "accession_no" : "BS0001" } }
{ "_id" : ObjectId("5c81fdd89cff5f2a54951b53"), "status_info" : { "accession_no" : "BS0002" } }
{ "_id" : ObjectId("5c81fe5e9cff5f2a54951b54"), "status_info" : { "accession_no" : "BS0003" } }
{ "_id" : ObjectId("5c81feaf9cff5f2a54951b55"), "status_info" : { "accession_no" : "BS0004" } }
{ "_id" : ObjectId("5c82005f9cff5f2a54951b56"), "status_info" : [ { "accession_no" : "BS0005" }, { "accession_no" : "BS0
006" }, { }, { "accession_no" : "BS0008" } ] }
{ "_id" : ObjectId("5c8201d69cff5f2a54951b57"), "status_info" : { "accession_no" : "BS0009" } }
{ "_id" : ObjectId("5c8202f69cff5f2a54951b58"), "status_info" : [ { "accession_no" : "BS010" }, { "accession_no" : "BS01
1" } ] }
```

To print records with accession no "BS0001"

```
> db.book_info.find({'status_info.accession_no':'BS0001'})
{ "_id" : ObjectId("5c81fd569cff5f2a54951b52"), "title" : "PROGRAMMING WITH JAVA", "status_info" : { "accession_no" : "B
S0001", "status" : "ISSUED" }, "author" : "EBALAGURUSWAMY", "cost" : "350", "publisher" : { "name" : "TMH", "location" :
 "banglore" } }
```

To print data in JSON format

```
Administrator: Command Prompt - mongo
> db.book_info.find().forEach(printjson)
{
        "_id" : ObjectId("5c81fd569cff5f2a54951b52"),
        "title" : "PROGRAMMING WITH JAVA",
        "status_info" : {
                "accession_no" : "BS0001",
                "status" : "ISSUED"
        },
        "author" : "EBALAGURUSWAMY",
        "cost" : "350",
        "publisher" : {
                "name" : "TMH",
                "location" : "banglore"
        }
}
{
        "_id" : ObjectId("5c81fdd89cff5f2a54951b53"),
        "title" : "ASP.NET3.5 VB 2008",
        "status_info" : {
                "accession_no" : "BS0002",
                "status" : "AVAIL"
        },
        "author" : "ANNE BOEHM",
        "cost" : "650",
        "publisher" : {
                "name" : "MURACH",
                "location" : "JAIPUR"
        }
}
{
```

To print the details of book where cost is more than 500

```
> db.book_info.find({"cost":{$gt:"500"}}).pretty()
{
        "_id" : ObjectId("5c81fdd89cff5f2a54951b53"),
        "title" : "ASP.NET3.5 VB 2008",
        "status_info" : {
                "accession_no" : "BS0002",
                "status" : "AVAIL"
        },
        "author" : "ANNE BOEHM",
        "cost" : "650",
        "publisher" : {
                "name" : "MURACH",
                "location" : "JAIPUR"
        }
}
{
        "_id" : ObjectId("5c81fe5e9cff5f2a54951b54"),
        "title" : "PRPGRAMMING IN VB",
        "status_info" : {
                "accession_no" : "BS0003",
                "status" : "ISSUED"
        },
        "author" : "JULIA CASE BRADLEY",
        "cost" : "600",
        "publisher" : {
                "name" : "TMH",
                "location" : "BANGLORE"
        }
```

Conclusion: We have successfully executed Data Collection, Data Curation and Management for large scale data system (such as MongoDB).