**VISVESVARAYA TECHNOLOGICAL UNIVERSITY**

**“JnanaSangama”, Belgaum -590014, Karnataka.**

****

**LAB REPORT**

**on**

**BIG DATA ANALYTICS**

**(20CS6PEBDA)**

***Submitted by***

**NAYANA J (1BM19CS095)**

***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 **NAYANA J(1BM19CS095),** 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 **Course Title - (Course code)**work prescribed for the said degree.

Nameof the Lab-Incharge               **Dr. Jyothi S Nayak**

Designation Professor and Head

Department of CSE Department of CSE

BMSCE, Bengaluru BMSCE, Bengaluru

`

**Index Sheet**

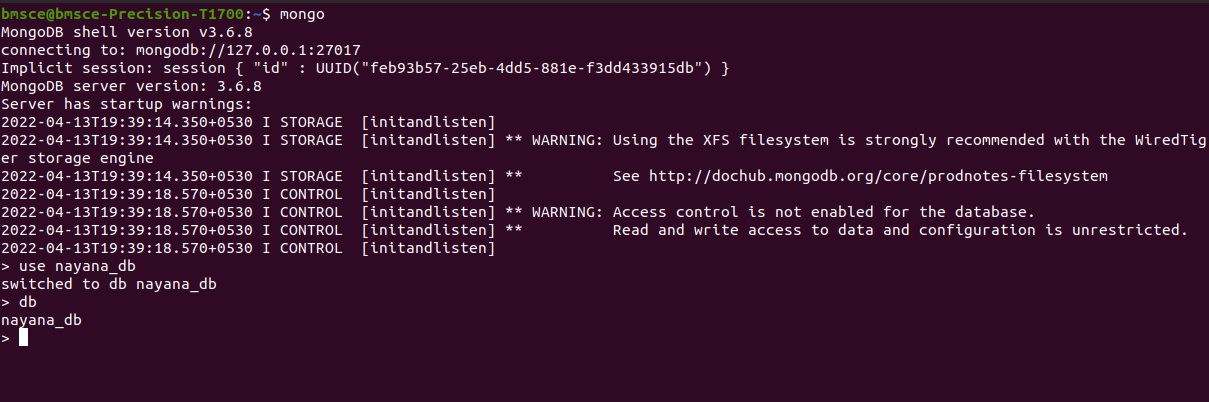
|  |  |  |
| --- | --- | --- |
| **Sl. No.** | **Experiment Title** | **Page No.** |
| **1.** | **MongoDB Lab Program 1 (CRUD Demonstration): -** Students should be classifying a dataset into one of the standard forms and apply suitable querying rules to obtain suitable results | **4** |
| **2.** | **MongoDB 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** |
| **3.** | **Cassandra Lab Program 1: -** 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) | **17** |
| **4.** | **Cassandra Lab Program 2: -** 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) | **19** |

**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 |

WEEK 1

1. **CREATE DATABASE IN MONGODB.**



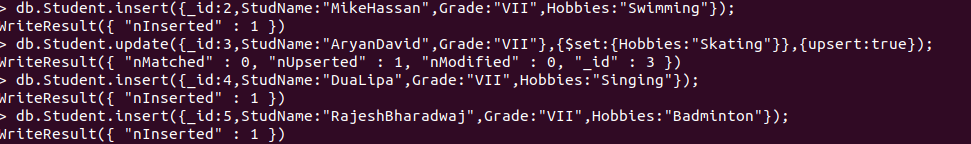
**2.** **CRUD (CREATE, READ, UPDATE, DELETE) OPERATIONS**

* To create a collection by the name “Student”. Let us take a look at the collection list prior to the creation of the new collection “Student”.



* Create a collection by the name “Students” and store the following data in it.



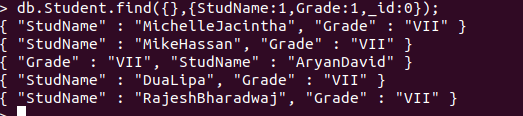


* FIND METHOD

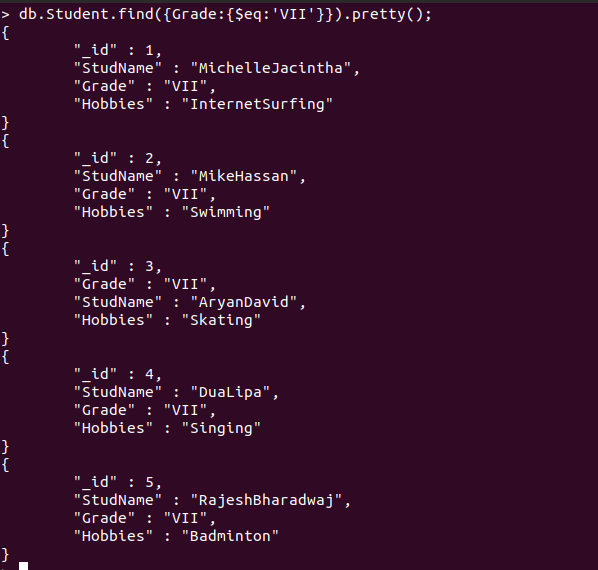
A. To search for documents from the “Students” collection based on certain search criteria.



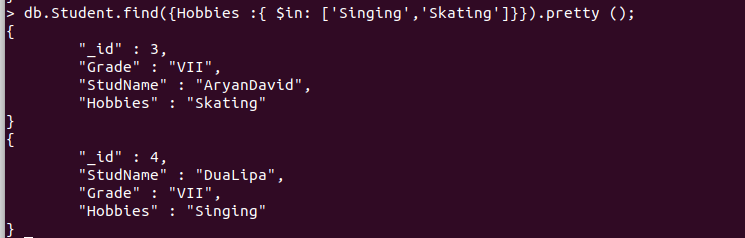
B. To display only the StudName and Grade from all the documents of the Students collection. The identifier\_id should be suppressed and NOT displayed.



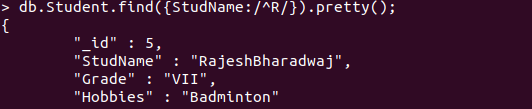
C. To find those documents where the Grade is set to ‘VII’



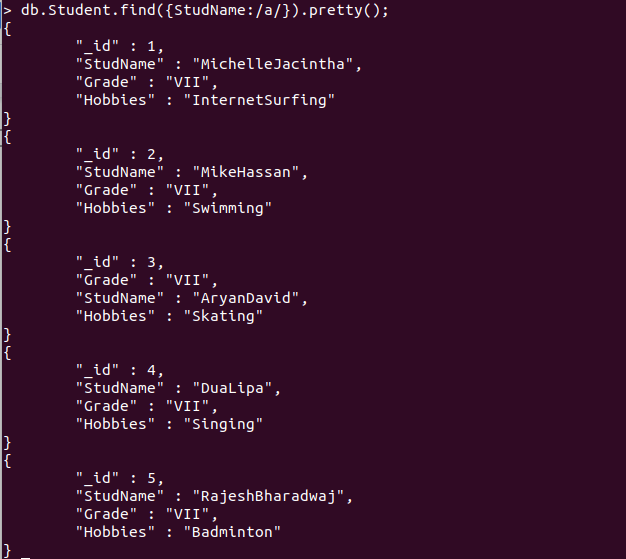
D. To find those documents from the Students collection where the Hobbies is set to either ‘singing’ or is set to ‘Skating’.



E. To find documents from the Students collection where the StudName begins with “R”.



F. To find documents from the Students collection where the StudNamehas an “a” in any position.



G. To find the number of documents in the Students collection.



H. To sort the documents from the Students collection in the descending order of StudName.

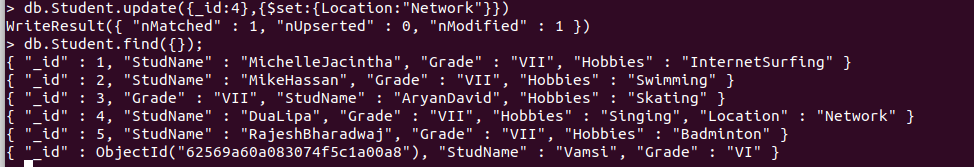


**3.****Save Method :**

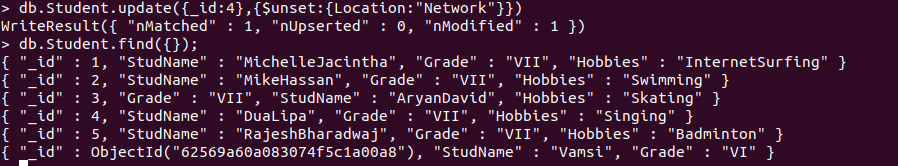
Save() method will insert a new document, if the document with the \_id does not exist. If it exists it will replace the exisiting document.



Add a new field to existing Document:



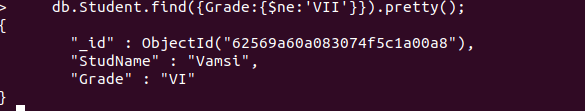
Remove the field in an existing Document



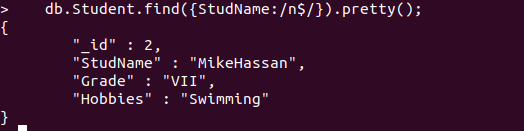
Finding Document based on search criteria suppressing few fields



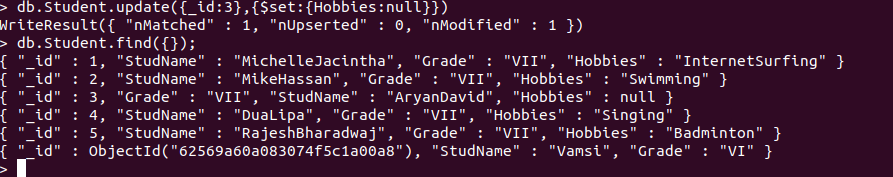
To find those documents where the Grade is not set to ‘VII’



To find documents from the Students collection where the StudName ends with n.



to set a particular field value to NULL



Count  the number of documents in Student Collections

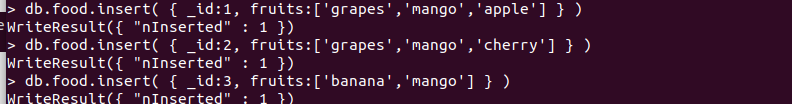


Count  the number  of documents in Student Collections with grade :VII



**food database using mongodb**

Create a collection by name “food” and add to each document add a “fruits” array



To find those documents from the “food” collection which has the “fruits array” constitute of “grapes”, “mango” and “apple”.

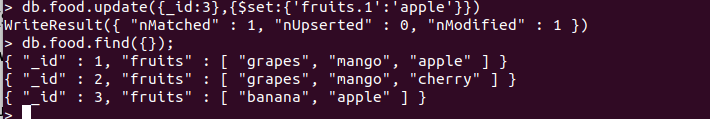


To find all the documets from the food collection which have elements mango and grapes in the array “fruits”



update on Array:

using particular id replace the element present in the 1st index position of the fruits array with apple

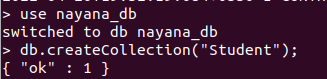


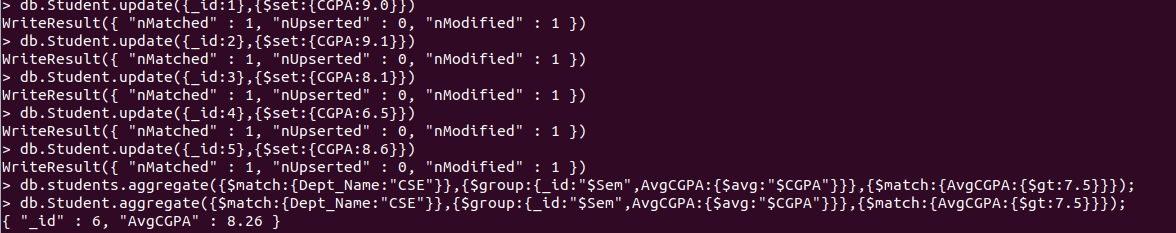
WEEK 2

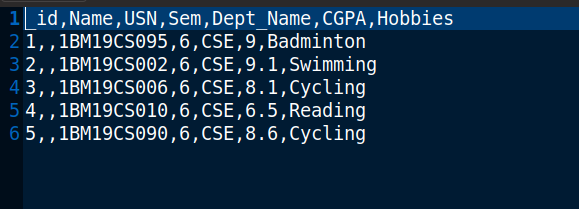
1. Using MongoDB
2. Create a database for Students and Create a Student Collection (\_id,Name, USN, Semester, Dept\_Name, CGPA, Hobbies(Set)).
3. Insert required documents to the collection.
4. First Filter on “Dept\_Name:CSE” and then group it on “Semester” and

compute the Average CPGA for that semester and flter those documents where the “Avg\_CPGA” is greater than 7.5.

1. Command used to export MongoDB JSON documents from “Student” Collection into the “Students” database into a CSV fle “Output.txt”.

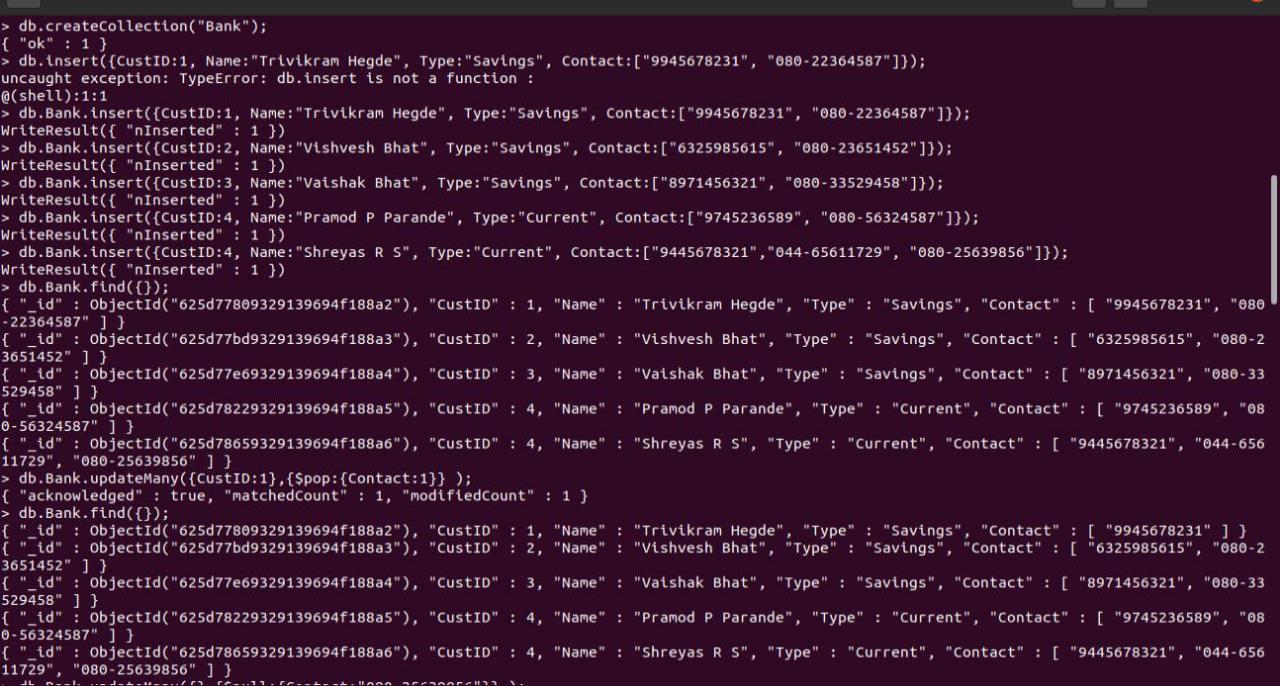


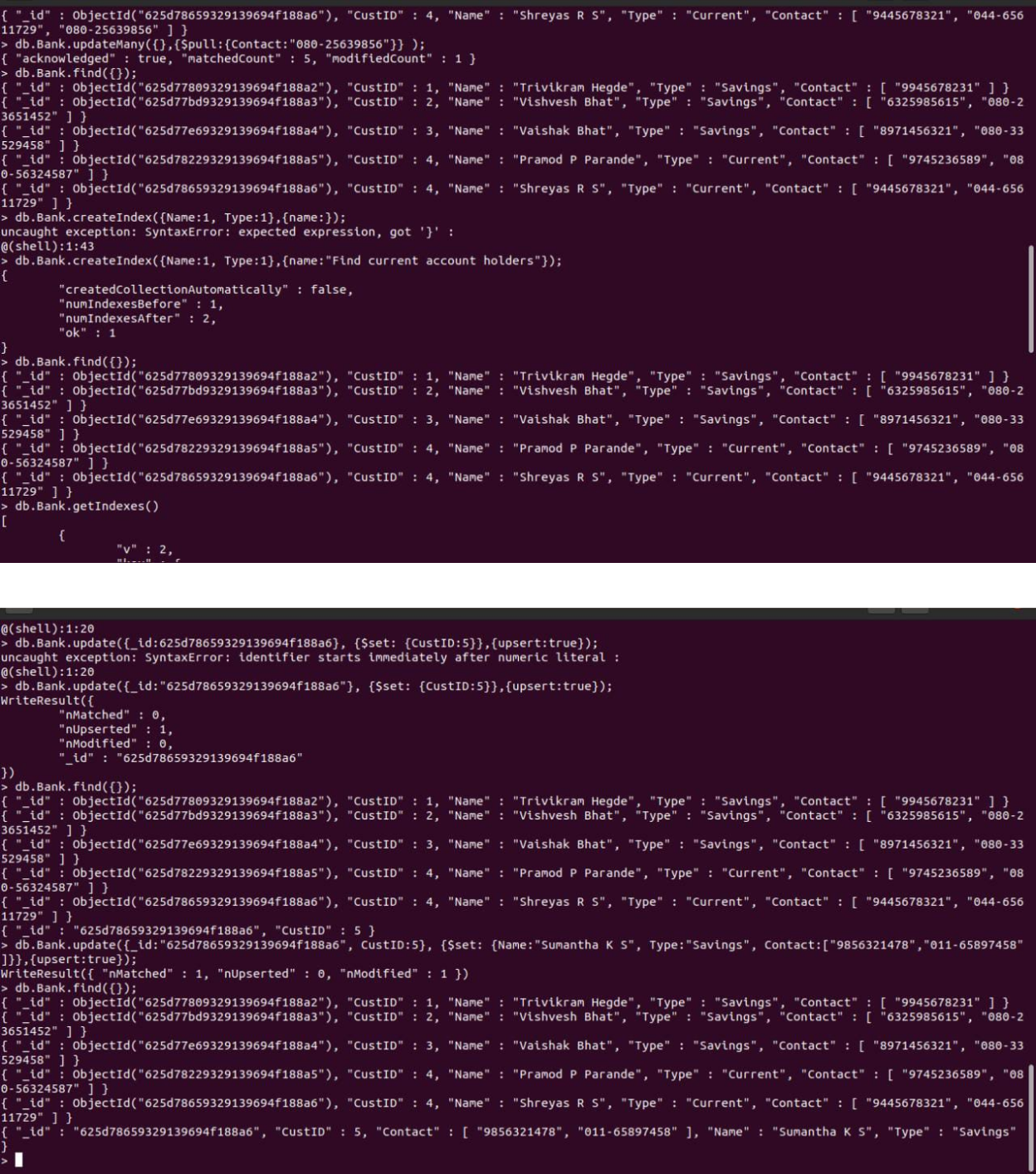




2)Create a mongodb collection Bank. Demonstrate the following by choosing felds of your choice.

1. Insert three documents
2. Use Arrays(Use Pull and Pop operation)
3. Use Index
4. Use Cursors
5. Updation





1) Using MongoDB,

1. Create a database for Faculty and Create a Faculty Collection(Faculty\_id, Name, Designation ,Department, Age, Salary, Specialization(Set)).
2. Insert required documents to the collection.
3. First Filter on “Dept\_Name:MECH” and then group it on “Designation” and

compute the Average Salary for that Designation and flter those documents where the “Avg\_Sal” is greater than 650000. iv) Demonstrate usage of import and export commands

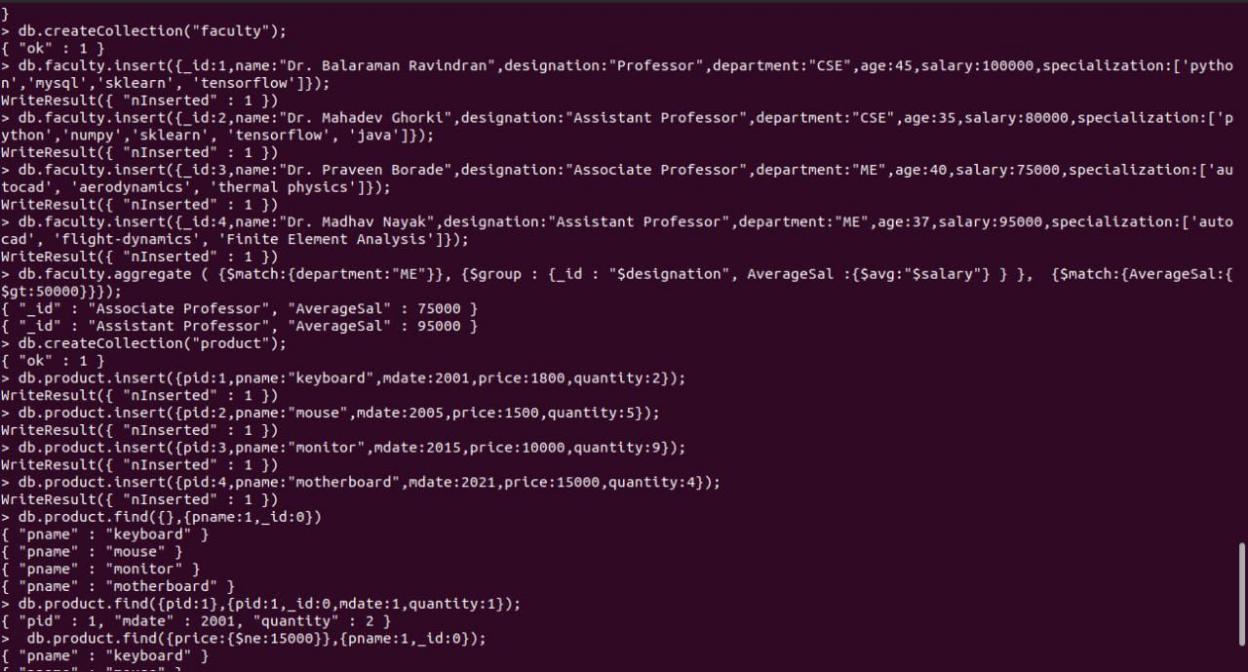
Write MongoDB queries for the following:

1)To display only the product name from all the documents of the product collection.

2)To display only the Product ID, ExpiryDate as well as the quantity from the document of the product collection where the \_id column is 1.

3)To fnd those documents where the price is not set to 15000.

4)To fnd those documents from the Product collection where the quantity is set to 9 and the product name is set to ‘monitor’.

5)To fnd documents from the Product collection where the Product name ends in ‘d’.

|  |  |  |
| --- | --- | --- |
| 3)Create a mongodb collection Hospital. Demonstrate the following by choosing felds of | | your |
| choice. | |  |
| 1 | Insert three documents |  |
| 2 | Use Arrays(Use Pull and Pop operation) |  |
| 3 | Use Index |  |
| 4 | Use Cursors |  |
| 5 | Updation |  |



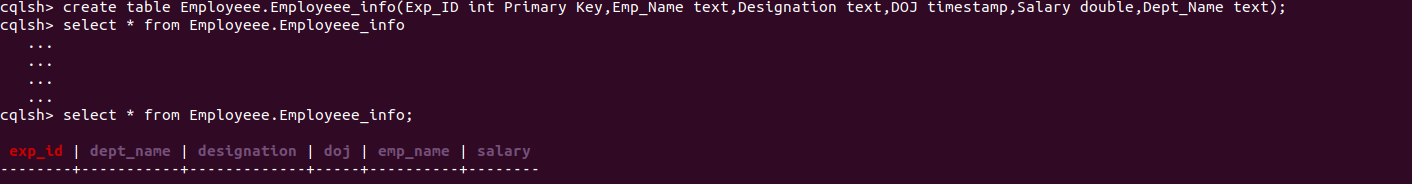
WEEK 3

Program 1. Perform the following DB operations using Cassandra.

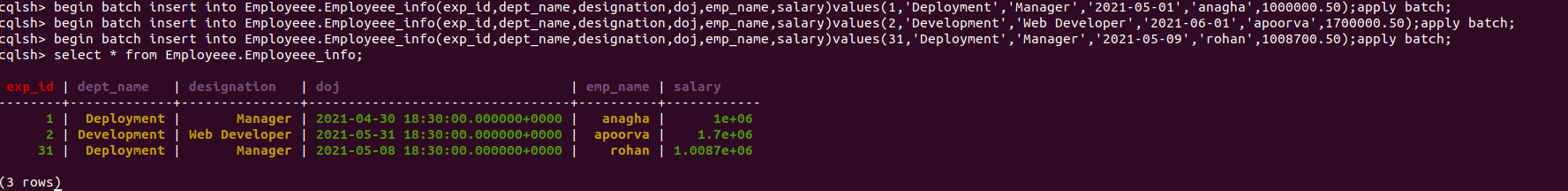
1. Create a key space by name Employee



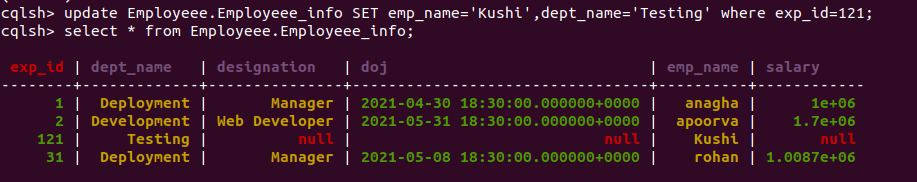
2. Create a column family by name Employee-Info with attributes Emp\_Id Primary Key, Emp\_Name, Designation, Date\_of\_Joining, Salary, Dept\_Name



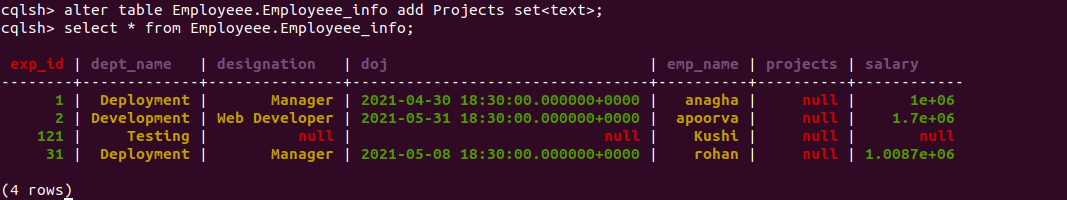
3. Insert the values into the table in batch



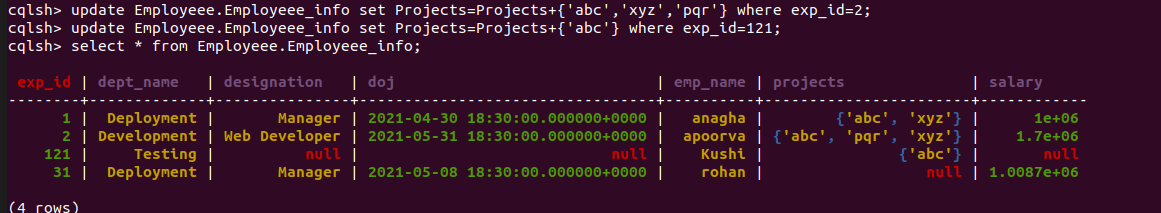
4. Update Employee name and Department of Emp-Id 121



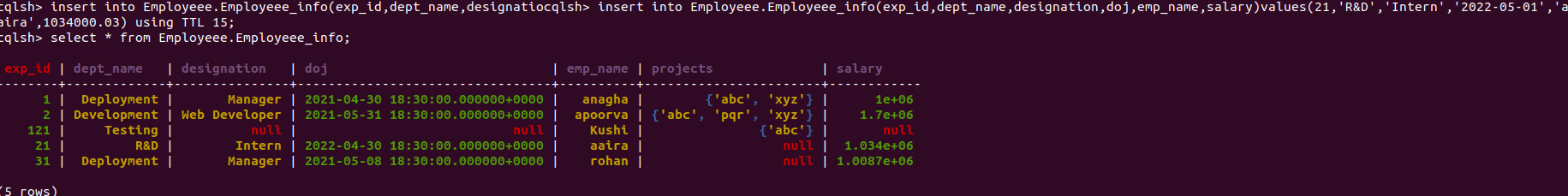
6.Alter the schema of the table Employee\_Info to add a column Projects which stores a set of Projects done by the corresponding Employee.

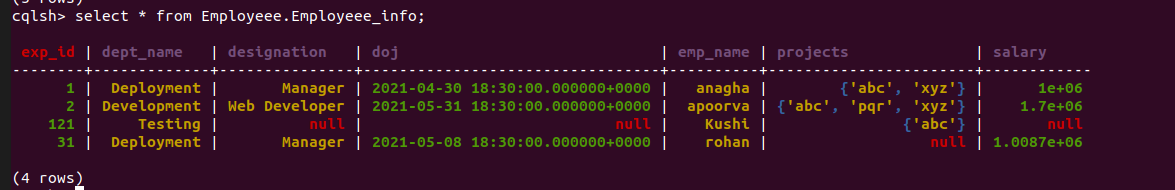


7. Update the altered table to add project names.



8 Create a TTL of 15 seconds to display the values of Employees.

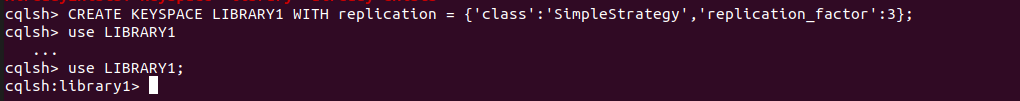




WEEK 4

**Program 2:**

**1 Create a key space by name Library**

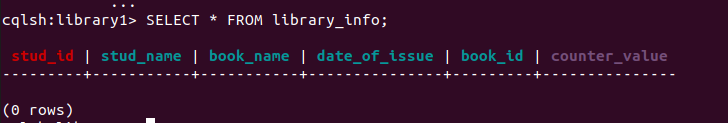
****

**2.Create a column family by name Library-Info with attributes**

**Stud\_Id Primary Key,**

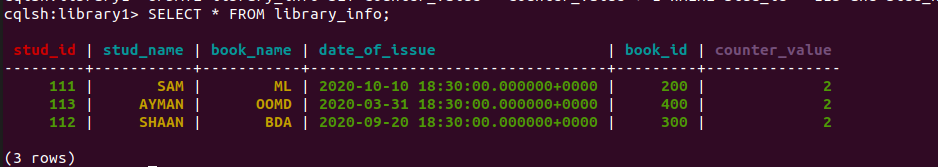
**Counter\_value of type Counter,**

**Stud\_Name, Book-Name, Book-Id, Date\_of\_issue**

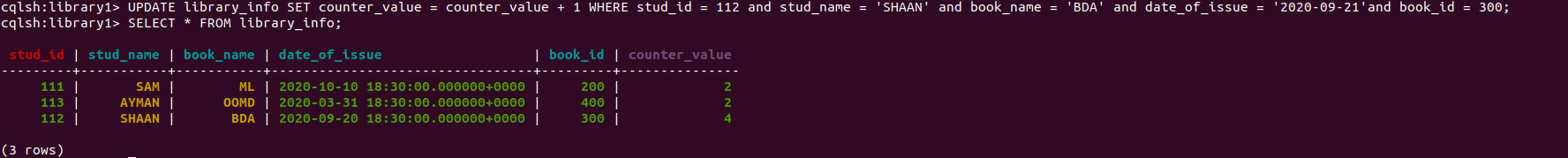
****

****

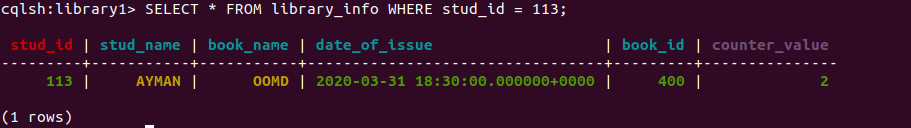
**3.Insert the values into the table in batch**

****

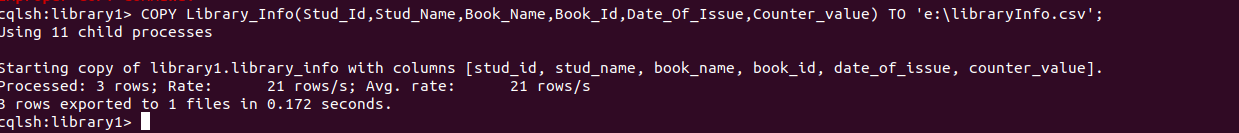
**4. Display the details of the table created and increase the value of the counter**

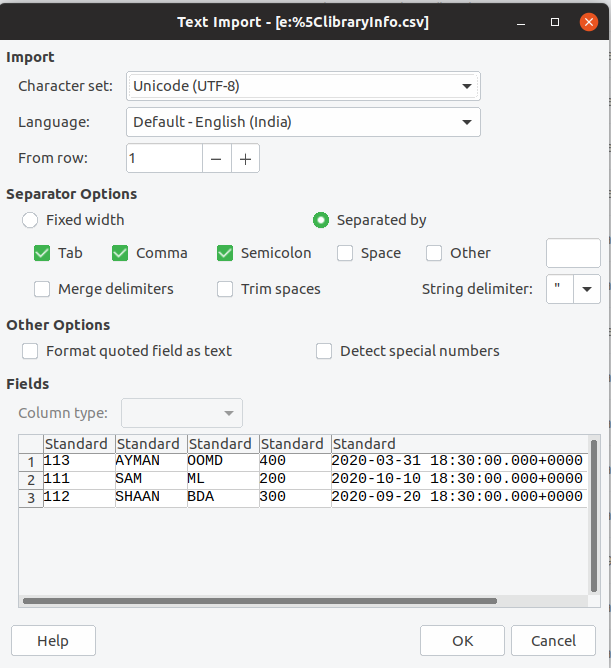
****

**5. Write a query to show that a student with id 113 has taken a book “OOMD” 2 times.**

****

**6. Export the created column to a csv file**

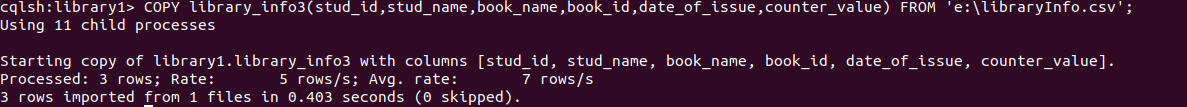
****

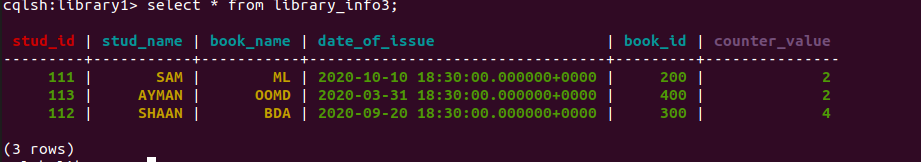
****

**7. Import a given csv dataset from local file system into Cassandra column**

**family**

****

****

****