

VISVESVARAYA TECHNOLOGICAL UNIVERSITY

“JnanaSangama”, Belgaum -590014, Karnataka.



LAB REPORT on

Big Data Analytics(23CS6PCBDA)

Submitted by

Maria Sayeema(1BM22CS151)

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

Feb-2024 to July-2024

**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 **Maria Sayeema(1BM22CS151)**, 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 2024. The Lab report has been approved as it satisfies the academic requirements in respect of a **Course Big Data Analytics(23CS6PCBDA)** work prescribed for the said degree.

Prof.Amrutha B
Assistant Professor
Department of CSE
BMSCE, Bengaluru

Dr. Kavitha Sooda
Professor and Head
Department of CSE
BMSCE, Bengaluru

Index Sheet

Sl. No.	Experiment Title	Page No.
1	MongoDB- CRUD Operations Demonstration (Practice and Self Study)	4
2	Neo4j	22
3	<p>Perform the following DB operations using Cassandra.</p> <ul style="list-style-type: none"> a) Create a keyspace by name Employee b) Create a column family by name Employee-Info with attributes Emp_Id Primary Key, Emp_Name, Designation, Date_of_Joining, Salary, Dept_Name c) Insert the values into the table in batch d) Update Employee name and Department of Emp-Id 121 e) Sort the details of Employee records based on salary f) Alter the schema of the table Employee_Info to add a column Projects which stores a set of Projects done by the corresponding Employee. g) Update the altered table to add project names. h) Create a TTL of 15 seconds to display the values of Employees. 	27
4	<p>Perform the following DB operations using Cassandra.</p> <ul style="list-style-type: none"> a) Create a keyspace by name Library b) 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 c) Insert the values into the table in batch d) Display the details of the table created and increase the value of the counter e) Write a query to show that a student with id 112 has taken a book "BDA" 2 times. f) Export the created column to a csv file 	30

	g) Import a given csv dataset from local file system into Cassandra column family	
5	Execution of HDFS Commands for interaction with Hadoop Environment. (Minimum 10 commands to be executed)	37
6	Implement Wordcount program on Hadoop framework	41
7	From the following link extract the weather data https://github.com/tomwhite/hadoop-book/tree/master/input/ncdc/all a) Create a MapReduce program to find average temperature for each year from NCDC data set. b) find the mean max temperature for every month.	48
8	Write a Scala program to print numbers from 1 to 100 using for loop.	55
9	Using RDD and FlatMap count how many times each word appears in a file and write out a list of words whose count is strictly greater than 4 using Spark.	57
10	Write a simple streaming program in Spark to receive text data streams on a particular port, perform basic text cleaning (like white space removal, stop words removal, lemmatization, etc.), and print the cleaned text on the screen. (Open Ended Question).	58

Course Outcome

CO1	Apply the concept of NoSQL, Hadoop or Spark for a given task
CO2	Analyse big data analytics mechanisms that can be applied to obtain solution for a given problem.
CO3	Design and implement solutions using data analytics mechanisms for a given problem.

Experiment - 1

MongoDB- CRUD Operations Demonstration

```
Activities Terminal Mar 11 14:30
mongosh mongodb+srv://<credentials>@cluster0.tud3r.mongodb.net/
Current Mongosh Log ID: 67cfffb58790ab5f3c1c0e7d
Connecting to: mongodb+srv://<credentials>@cluster0.tud3r.mongodb.net/?appName=mongosh+2.2.1

Using MongoDB: 8.0.5 (API Version 1)
Using Mongosh: 2.2.1
mongosh 2.4.2 is available for download: https://www.mongodb.com/try/download/shell

For mongosh info see: https://docs.mongodb.com/mongodb-shell/

Atlas atlas-vvj19l-shard-0 [primary] test>
Atlas atlas-vvj19l-shard-0 [primary] test> use myDB
switched to db myDB
Atlas atlas-vvj19l-shard-0 [primary] myDB> create.collection("Student1");
ReferenceError: create is not defined
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.createCollection("Student1");
Uncought:
SyntaxError: Unexpected token, expected "," (1:29)
> 1 | db.createCollection("Student1");
   |
   2 |

Atlas atlas-vvj19l-shard-0 [primary] myDB> db.createCollection("Student1");
ok: 1
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student1.insert({
...   _id: 1,
...   StudName: 'MichelleJacintha',
...   Grade: 'VII',
...   Hobbies: 'InternetSurfing'
... });
DeprecationWarning: collection.insert() is deprecated. Use insertOne, insertMany, or bulkWrite.
[ acknowledged: true, insertedIds: { '_id': 1 } ]
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student1.update(
...   { _id: 3, StudName: 'Arvandavid' }, Grade: 'VII',
...   { $set: { Hobbies: 'Skating' } },
...   { upsert: true }
... );
DeprecationWarning: collection.update() is deprecated. Use updateOne, updateMany, or bulkWrite.
{
  acknowledged: true,
  insertedIds: 1,
  matchedCount: 0,
  modifiedCount: 0,
  upsertedCount: 1
}
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student1.find();
[{
  _id: 1,
  StudName: 'MichelleJacintha',
  Grade: 'VII',
  Hobbies: 'InternetSurfing'
},
{ _id: 3, Grade: 'VII', StudName: 'Arvandavid', Hobbies: 'Skating' }]
```

```

Activities Terminal Mar 11 14:31
mongosh mongoDB+srv://<credentials>@cluster0.tud3r.mongodb.net/
[{"_id": 3, "Grade": "VII", "StudName": "AryanDavid", "Hobbies": "Skating"}]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.update(
...   {"_id": 3, "StudName": "AryanDavid", "Grade": "VII"}, 
...   {"$set": {"Hobbies": "Chess"}}, 
... );
{
  acknowledged: true,
  insertedCount: 0,
  matchedCount: 1,
  modifiedCount: 1,
  upsertedCount: 0
}
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find();
ReferenceError: ds is not defined
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find();
[
  {
    "_id": 1,
    "StudName": "MichelleAcintha",
    "Grade": "VII",
    "Hobbies": "InternetSurfing"
  },
  {"_id": 3, "Grade": "VII", "StudName": "AryanDavid", "Hobbies": "Chess"}
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({StudName: "Aryan David"});
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({StudName: "Aryan Davld"});
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({StudName: "Aryandavid"});
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({StudName: "MichelleAcintha", Grade: "VII"}, {Grade: "VII", StudName: "Aryandavid"})
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({StudName: "MichelleAcintha", Grade: "VII"}, {Grade: "VII", StudName: "AryanDavid"});
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({Grade: {"$eq": "VII"}}).pretty();
[
  {
    "_id": 3,
    "StudName": "MichelleAcintha",
    "Grade": "VII",
    "Hobbies": "InternetSurfing"
  },
  {"_id": 3, "Grade": "VII", "studName": "AryanDavid", "Hobbies": "Chess"}
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({hobbies: { $in: [ "Chess", "Skating" ] }}).pretty();
[{"_id": 3, "Grade": "VII", "studName": "AryanDavid", "Hobbies": "Chess"}]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({studName: /M/}).pretty();
[{"_id": 1, "StudName": "MichelleAcintha", "Grade": "VII", "Hobbies": "InternetSurfing"}]

Activities Terminal Mar 11 14:32
mongosh mongoDB+srv://<credentials>@cluster0.tud3r.mongodb.net/
[{"_id": 3, "Grade": "VII", "StudName": "AryanDavid", "Hobbies": "Chess"}]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({}, {StudName: 1, GrGrade: 1, _id: 0});
[
  {
    "StudName": "MichelleAcintha",
    "Grade": "VII"
  },
  {"Grade": "VII", "StudName": "Aryandavid"}
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({}, {StudName: 1, Grade: 1, _id: 0});
[
  {
    "StudName": "MichelleAcintha",
    "Grade": "VII"
  },
  {"Grade": "VII", "StudName": "Aryandavid"}
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({Grade: {"$eq": "VII"}}).pretty();
[
  {
    "_id": 1,
    "StudName": "MichelleAcintha",
    "Grade": "VII",
    "Hobbies": "InternetSurfing"
  },
  {"_id": 3, "Grade": "VII", "StudName": "AryanDavid", "Hobbies": "Chess"}
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({hobbies: { $in: [ "Chess", "Skating" ] }}).pretty();
[{"_id": 3, "Grade": "VII", "studName": "AryanDavid", "Hobbies": "Chess"}]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({studName: /e/}).pretty();
[{"_id": 1, "StudName": "MichelleAcintha", "Grade": "VII", "Hobbies": "InternetSurfing"}]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.count();
DeprecationWarning: Collection.count() is deprecated. Use countDocuments or estimatedDocumentCount.
2
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find().sort({StudName: -1}).pretty();
[
  {
    "_id": 1,
    "StudName": "MichelleAcintha",
    "Grade": "VII",
    "Hobbies": "InternetSurfing"
  },
  {"_id": 3, "Grade": "VII", "StudName": "AryanDavid", "Hobbies": "Chess"}
]
Atlas atlas-vvji9t-shard-0 [primary] myDB>

```

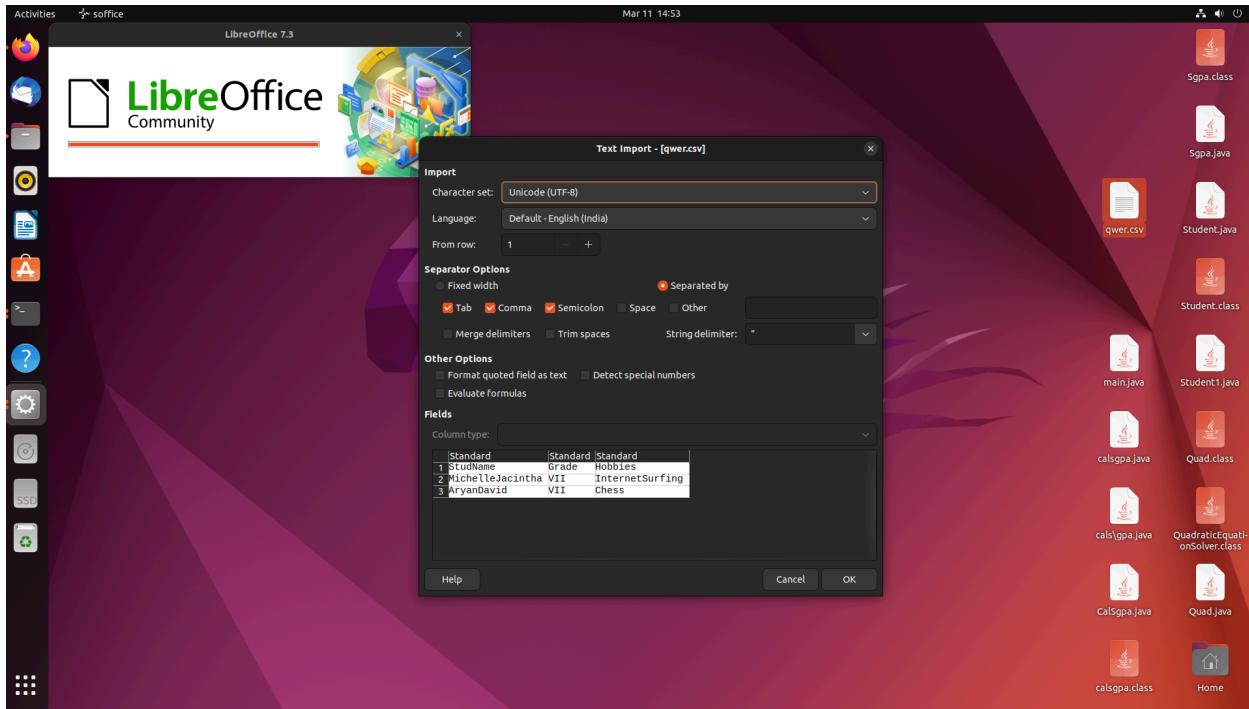
```
Activities Terminal Mar 11 14:31
mongosh mongodb+srv://[credentials]@cluster0.tud3r.mongodb.net/
{
  "_id": 3, Grade: 'VII', StudName: 'AryanDavid', Hobbies: 'Skating' }
]
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student1.update(
...  { _id: 3, StudName: 'AryanDavid', Grade: 'VII' },
...  { $set: {hobbies: 'Chess' } }
... );
{
  acknowledged: true,
  insertedId: null,
  matchedCount: 1,
  modifiedCount: 1,
  upsertedCount: 0
}
Atlas atlas-vvj19l-shard-0 [primary] myDB> ds.Student1.find();
ReferenceError: ds is not defined
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student1.find();
[
  {
    _id: 1,
    StudName: 'MichelleJacintha',
    Grade: 'VII',
    Hobbies: 'InternetSurfing'
  },
  {
    _id: 3, Grade: 'VII', StudName: 'AryanDavid', Hobbies: 'Chess' }
]
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student.find({StudName: "Aryan David"});
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student.find({StudName: "Aryan David"});
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student1.find({StudName: "AryanDavid"});
[ { _id: 3, Grade: 'VII', StudName: 'AryanDavid', Hobbies: 'Chess' } ]
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student1.find({}, {Studname: 1, GrGrade: 1, _id: 0});
[
  { StudName: 'MichelleJacintha', Grade: 'VII' },
  { Grade: 'VII', StudName: 'AryanDavid' }
]
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student1.find({}, {StudName: 1, Grade: 1, _id: 0});
[ { StudName: 'MichelleJacintha', Grade: 'VII' },
  { Grade: 'VII', StudName: 'AryanDavid' } ]
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student1.find({Grade: {seq: 'VII'}}).pretty();
[
  {
    _id: 1,
    StudName: 'MichelleJacintha',
    Grade: 'VII',
    Hobbies: 'Internetsurfing'
  },
  {
    _id: 3, Grade: 'VII', StudName: 'AryanDavid', Hobbies: 'Chess' }
]
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student1.find({Hobbies: { $in: [ 'Chess', 'Skating' ] }}).pretty();
[ { _id: 3, Grade: 'VII', StudName: 'AryanDavid', Hobbies: 'Chess' } ]
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student1.find({StudName: /A/}).pretty();
[
```

```
Activities Terminal Mar 11 14:32
mongosh mongodb+srv://[credentials]>@cluster0.tud3r.mongodb.net/
[ { _id: 3, Grade: 'VII', StudName: 'AryanDavid', Hobbies: 'Chess' } ]
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student1.find({}, {studName: 1, GrGrade: 1, _id: 0});
[
  { StudName: 'MichelleJacintha', Grade: 'VII' },
  { Grade: 'VII', StudName: 'AryanDavid' }
]
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student1.find({}, {studName: 1, Grade: 1, _id: 0});
[
  { StudName: 'MichelleJacintha', Grade: 'VII' },
  { Grade: 'VII', StudName: 'AryanDavid' }
]
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student1.find({Grade: { $eq: 'VII'}}).pretty();
[
  {
    "_id": 1,
    "StudName": "MichelleJacintha",
    "Grade": "VII",
    "Hobbies": "Internetsurfing"
  },
  {
    "_id": 3, Grade: 'VII', StudName: 'AryanDavid', Hobbies: 'Chess' }
]
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student1.find({Hobbies: { $in: [ 'Chess', 'Skating' ] }}).pretty();
[
  { _id: 3, Grade: 'VII', StudName: 'AryanDavid', Hobbies: 'Chess' }
]
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student1.find({StudName: /A/}).pretty();
[
  {
    "_id": 1,
    "StudName": "MichelleJacintha",
    "Grade": "VII",
    "Hobbies": "Internetsurfing"
  }
]
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student1.find({studName: /e/}).pretty();
[
  {
    "_id": 1,
    "StudName": "MichelleJacintha",
    "Grade": "VII",
    "Hobbies": "Internetsurfing"
  }
]
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student1.count();
DeprecationWarning: Collection.count() is deprecated. Use countDocuments or estimatedDocumentCount.
1
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student1.find().sort({studName: -1}).pretty();
[
  {
    "_id": 1,
    "StudName": "MichelleJacintha",
    "Grade": "VII",
    "Hobbies": "Internetsurfing"
  },
  { _id: 3, Grade: 'VII', StudName: 'AryanDavid', Hobbies: 'Chess' }
]
Atlas atlas-vvj19l-shard-0 [primary] myDB>
```

```

Activities Terminal Mar 11 14:52
bmseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: ~
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoexport
2025-03-11T14:39:05.282+0530  must specify a collection
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ try "mongoexport --help" for more information
mongodump command not found
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ ^[[200-mongosh "mongodb+srv://cluster0.tud3r.mongodb.net/" --apiVersion 1 --username mariacs22 --password 8Ja9u0zIrEjzBlot
-mongosh: command not found
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ Current Mongosh Log ID: 67cffdfcfa3c80154c00e7d
Connecting to: mongodb+srv://credentials@cluster0.tud3r.mongodb.net/?appName=mongosh+2.2.1
Using MongoDB: 8.0.5 (API Version 1)
Using Mongosh: 2.4.2
mongosh 2.4.2 ls is available for download: https://www.mongodb.com/try/download/shell
For mongosh info see: https://docs.mongodb.com/mongodb-shell/
Atlas atlas-vvji9l-shard-0 [primary] test> use myDB;
switched to db myDB
Atlas atlas-vvji9l-shard-0 [primary] myDB> mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "Stu
dName,Grade,Hobbies"
Unccaught:
SyntaxError: Missing semicolon. (1:14)
> 1 | mongoexport -uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "StuName,Grade,Hobbies"
2 |
Atlas atlas-vvji9l-shard-0 [primary] myDB> mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "Stu
dName,Grade,Hobbies"
Unccaught:
SyntaxError: Missing semicolon. (1:14)
> 1 | mongoexport -uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "StuName,Grade,Hobbies"
2 |
Atlas atlas-vvji9l-shard-0 [primary] myDB> (To exit, press Ctrl+C again or Ctrl+D or type .exit)
Atlas atlas-vvji9l-shard-0 [primary] myDB> mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "Stu
dName,Grade,Hobbies"
Unccaught:
SyntaxError: Missing semicolon. (1:14)
> 1 | mongoexport -uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "StuName,Grade,Hobbies"
2 |
Atlas atlas-vvji9l-shard-0 [primary] myDB> exit
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --
fields "StuName,Grade,Hobbies"
2025-03-11T14:42:20.815+0530 connected to: mongodb+srv://[*REDACTED*]@cluster0.tud3r.mongodb.net/myDB
2025-03-11T14:42:21.068+0530 exported 2 records
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/out.txt" --file
ids "StuName,Grade,Hobbies"
Activities Terminal Mar 11 14:53
bmseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: ~
dnName,Grade,Hobbies"
Unccaught:
SyntaxError: Missing semicolon. (1:14)
> 1 | mongoexport -uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "StuName,Grade,Hobbies"
2 |
Atlas atlas-vvji9l-shard-0 [primary] myDB> (To exit, press Ctrl+C again or Ctrl+D or type .exit)
Atlas atlas-vvji9l-shard-0 [primary] myDB> mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "Stu
dName,Grade,Hobbies"
Unccaught:
SyntaxError: Missing semicolon. (1:14)
> 1 | mongoexport -uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "StuName,Grade,Hobbies"
2 |
Atlas atlas-vvji9l-shard-0 [primary] myDB> exit
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --
fields "StuName,Grade,Hobbies"
2025-03-11T14:42:20.815+0530 connected to: mongodb+srv://[*REDACTED*]@cluster0.tud3r.mongodb.net/myDB
2025-03-11T14:42:21.068+0530 exported 2 records
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/out.txt" --file
ids "StuName,Grade,Hobbies"
2025-03-11T14:43:35.174+0530 connected_to: mongodb+srv://[*REDACTED*]@cluster0.tud3r.mongodb.net/myDB
2025-03-11T14:43:35.174+0530 exported 2 records
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/qwer.txt" --file
ids "StuName,Grade,Hobbies"
2025-03-11T14:44:22.053+0530 connected_to: mongodb+srv://[*REDACTED*]@cluster0.tud3r.mongodb.net/myDB
2025-03-11T14:44:22.053+0530 exported 2 records
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ whoami
bnseccse
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ ^C
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "/home/bmseccse/De
sktop/qwer.csv"
2025-03-11T14:47:53.277+0530 connected_to: mongodb+srv://[*REDACTED*]@cluster0.tud3r.mongodb.net/myDB
2025-03-11T14:47:53.277+0530 exported 2 records
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoimport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student3 --type csv --file "/home/bmseccse/D
esktop/qwer.csv" --headerline
2025-03-11T14:51:32.105+0530 connected_to: mongodb+srv://[*REDACTED*]@cluster0.tud3r.mongodb.net/myDB
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ 

```



```
Activities Terminal Mar 11 15:21
mongosh mongodb+srv://<credentials>@cluster0.tud3r.mongodb.net/
Current Mongosh log ID: 67cfffb58790ab5f31c00e7d
Connecting to: mongodb+srv://<credentials>@cluster0.tud3r.mongodb.net/?appName=mongosh+2.2.1

Using MongoDB: 8.0.5 (API Version 1)
Using Mongosh: 2.2.1
mongosh 2.4.2 is available for download: https://www.mongodb.com/try/download/shell

For mongosh info see: https://docs.mongodb.com/mongodb-shell/

Atlas atlas-vvj19l-shard-0 [primary] test>
Atlas atlas-vvj19l-shard-0 [primary] test> use myDB
switched to db myDB
Atlas atlas-vvj19l-shard-0 [primary] myDB> create.collection("Student");
ReferenceError: create is not defined
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.createCollection("Student");
Unccaught:
SyntaxError: Unexpected token, expected "," (1:29)
> 1 | db.createCollection("Student");
   |
2 |
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.createCollection("Student");
{ ok: 1 }
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student1.insert({
   ...
   _id: 1,
   ...
   StudName: "MichelleJacintha",
   ...
   Grade: "VII",
   ...
   Hobbies: "Internetsurfing"
   ...
 })
DeprecationWarning: Collection.insert() is deprecated. Use insertOne, insertMany, or bulkWrite.
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student1.update(
   ...
   { _id: 1, Studname: "AryanDavid" },
   ...
   { $set: {Hobbies: "Skating"} },
   ...
   { upsert: true }
   ...
 )
DeprecationWarning: Collection.update() is deprecated. Use updateOne, updateMany, or bulkWrite.
{
   acknowledged: true,
   insertedId: 3,
   matchedCount: 0,
   modifiedCount: 0,
   upsertedCount: 1
}
Atlas atlas-vvj19l-shard-0 [primary] myDB> db.Student1.find();
[[
   {
      _id: 1,
      StudName: 'MichelleJacintha',
      Grade: 'VII',
      Hobbies: 'Internetsurfing'
   },
   {
      _id: 3, Grade: 'VII', StudName: 'AryanDavid', Hobbies: 'Skating'
   }
]]
```

```
Activities Terminal Mar 11 15:22
mongosh mongodb+srv://<credentials>@cluster0.tud3r.mongodb.net/
[{"_id": 3, "Grade": "VII", "StudName": "AryanDavid", "Hobbies": "Skating"}]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.update(
...   {"_id": 3, "StudName": "AryanDavid", "Grade": "VII"}, 
...   {"$set": {"Hobbies": "Chess"}}
... );
{
  "acknowledged": true,
  "insertedId": null,
  "matchedCount": 1,
  "modifiedCount": 1,
  "upsertedCount": 0
}
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find();
ReferenceError: ds is not defined
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find();
[
  {
    "_id": 3,
    "StudName": "MichelleJacintha",
    "Grade": "VII",
    "Hobbies": "InternetSurfing"
  },
  {"_id": 3, "Grade": "VII", "StudName": "AryanDavid", "Hobbies": "Chess"}
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({StudName: "Aryan David"});
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({StudName: "AryanDavid"});
[{"_id": 3, "Grade": "VII", "StudName": "AryanDavid", "Hobbies": "Chess"}]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({}, {StudName: 1, _id: 0});
[
  {"StudName": "MichelleJacintha", "Grade": "VII"}, 
  {"Grade": "VII", "StudName": "AryanDavid"}
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({}, {StudName: 1, Grade: 1, _id: 0});
[
  {"StudName": "MichelleJacintha", "Grade": "VII"}, 
  {"Grade": "VII", "StudName": "AryanDavid"}
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({Grade: {"$eq": "VII"}}).pretty();
[
  {
    "_id": 3,
    "StudName": "MichelleJacintha",
    "Grade": "VII",
    "Hobbies": "InternetSurfing"
  },
  {"_id": 3, "Grade": "VII", "StudName": "AryanDavid", "Hobbies": "Chess"}
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({Hobbies: { $in: [ "Chess", "Skating" ] }}).pretty();
[{"_id": 3, "Grade": "VII", "StudName": "AryanDavid", "Hobbies": "Chess"}]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({StudName: /M/}).pretty();
[{"_id": 3, "Grade": "VII", "StudName": "AryanDavid", "Hobbies": "Chess"}]
```

```
Activities Terminal Mar 11 15:22
mongosh mongodb+srv://<credentials>@cluster0.tud3r.mongodb.net/
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({StudName: /A/}).pretty();
[
  {
    "_id": 3,
    "StudName": "MichelleJacintha",
    "Grade": "VII",
    "Hobbies": "InternetSurfing"
  }
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({StudName: /e/}).pretty();
[
  {
    "_id": 3,
    "StudName": "MichelleJacintha",
    "Grade": "VII",
    "Hobbies": "InternetSurfing"
  }
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.count();
DeprecationWarning: Collection.count() is deprecated. Use countDocuments or estimatedDocumentCount.
2
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find().sort({StudName: -1}).pretty();
[
  {
    "_id": 3,
    "StudName": "MichelleJacintha",
    "Grade": "VII",
    "Hobbies": "InternetSurfing"
  },
  {"_id": 3, "Grade": "VII", "StudName": "AryanDavid", "Hobbies": "Chess"}
]
Atlas atlas-vvji9t-shard-0 [primary] myDB>
(To exit, press Ctrl+c again or Ctrl+d or type .exit)
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find().pretty();
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student3.find();
[
  {
    "_id": ObjectId("67d0009c20453206f8c4ffcc"),
    "StudName": "AryanDavid",
    "Grade": "VII",
    "Hobbies": "Chess"
  },
  {
    "_id": ObjectId("67d0009c20453206f8c4ffcd"),
    "StudName": "MichelleJacintha",
    "Grade": "VII",
    "Hobbies": "InternetSurfing"
  }
]
```

```

Activities Terminal Mar 11 15:21
mongosh mongodb+srv://<credentials>@cluster0.tud3r.mongodb.net/
Current Mongosh Log ID: 67cff8b58796ab5f3c100e7d
Connecting to: mongodb+srv://<credentials>@cluster0.tud3r.mongodb.net/?appName=mongosh+2.2.1
Using MongoDB: 8.0.5 (API Version 1)
Using Mongosh: 2.2.1
mongosh 2.4.2 is available for download: https://www.mongodb.com/try/download/shell
For mongosh info see: https://docs.mongodb.com/mongodb-shell/
Atlas atlas-vvji9l-shard-0 [primary] test>
Atlas atlas-vvji9l-shard-0 [primary] test> use myDB
switched to db myDB
Atlas atlas-vvji9l-shard-0 [primary] myDB> create.collection("student1");
ReferenceError: create is not defined
Atlas atlas-vvji9l-shard-0 [primary] myDB> db.createCollection("Student1");
Uncaught:
SyntaxError: Unexpected token, expected "," (1:29)
> 1 | db.createCollection("Student1");
|   ^
| 2 |
Atlas atlas-vvji9l-shard-0 [primary] myDB> db.createCollection("Student1");
{ ok: 1 }
Atlas atlas-vvji9l-shard-0 [primary] myDB> db.Student1.insert({
... _id: 1,
... StudName: "MichelleJacintha",
... Grade: "VII",
... Hobbies: "Internetsurfing"
... });
DeprecationWarning: Collection.insert() is deprecated. Use insertOne, insertMany, or bulkWrite.
{
  acknowledged: true,
  insertedId: 1,
  matchedCount: 0,
  modifiedCount: 0,
  upsertedCount: 1
}
Atlas atlas-vvji9l-shard-0 [primary] myDB> db.Student1.update(
... { _id: 1, StudName: "AryanDavid", Grade: "VII" },
... { $set: { hobbies: "Skating" } },
... { upsert: true }
... );
DeprecationWarning: Collection.update() is deprecated. Use updateOne, updateMany, or bulkWrite.
{
  acknowledged: true,
  insertedId: 1,
  matchedCount: 0,
  modifiedCount: 0,
  upsertedCount: 1
}
Atlas atlas-vvji9l-shard-0 [primary] myDB> db.Student1.find();
[
  {
    _id: 1,
    StudName: "MichelleJacintha",
    Grade: "VII",
    Hobbies: "Internetsurfing"
  }
]
{ _id: 1, Grade: "VII", StudName: "AryanDavid", Hobbies: "Skating" }

Activities Terminal Mar 11 15:22
mongosh mongodb+srv://<credentials>@cluster0.tud3r.mongodb.net/
[ _id: 3, Grade: "VII", StudName: "AryanDavid", Hobbies: "Skating" ]
Atlas atlas-vvji9l-shard-0 [primary] myDB> db.Student1.update(
... { _id: 3, studname: "AryanDavid", Grade: "VII" },
... { $set: { hobbies: "Chess" } }
... );
{
  acknowledged: true,
  insertedId: null,
  matchedCount: 1,
  modifiedCount: 1,
  upsertedCount: 0
}
Atlas atlas-vvji9l-shard-0 [primary] myDB> ds.Student1.find();
[
  {
    _id: 1,
    StudName: "MichelleJacintha",
    Grade: "VII",
    Hobbies: "Internetsurfing"
  },
  {
    _id: 3, Grade: "VII", StudName: "AryanDavid", Hobbies: "Chess"
  }
]
Atlas atlas-vvji9l-shard-0 [primary] myDB> db.Student.find({StudName: "Aryan David"});
Atlas atlas-vvji9l-shard-0 [primary] myDB> db.Student.find({StudName: "Aryan David"});
Atlas atlas-vvji9l-shard-0 [primary] myDB> db.Student1.find({StudName: "AryanDavid"});
[ { _id: 3, Grade: "VII", StudName: "AryanDavid", Hobbies: "Chess" } ]
Atlas atlas-vvji9l-shard-0 [primary] myDB> db.Student1.find({}, {StudName: 1, GrGrade: 1, _id: 0});
[ { StudName: "MichelleJacintha", Grade: "VII" },
  { Grade: "VII", StudName: "AryanDavid" }
]
Atlas atlas-vvji9l-shard-0 [primary] myDB> db.Student1.find({}, {StudName: 1, Grade: 1, _id: 0});
[ { StudName: "MichelleJacintha", Grade: "VII" },
  { Grade: "VII", StudName: "AryanDavid" }
]
Atlas atlas-vvji9l-shard-0 [primary] myDB> db.Student1.find({Grade: {Seq: "VII"}}).pretty();
[
  {
    _id: 3,
    StudName: "MichelleJacintha",
    Grade: "VII",
    Hobbies: "Internetsurfing"
  },
  {
    _id: 3, Grade: "VII", StudName: "AryanDavid", Hobbies: "Chess"
  }
]
Atlas atlas-vvji9l-shard-0 [primary] myDB> db.Student1.find({hobbies: { $in: ['Chess', 'Skating'] }}).pretty();
[ { _id: 3, Grade: "VII", StudName: "AryanDavid", Hobbies: "Chess" } ]
Atlas atlas-vvji9l-shard-0 [primary] myDB> db.Student1.find({studName: /^M/}).pretty();
[ { }
]

```

```

Activities Terminal Mar 11 15:22
mongosh mongodb+srv://<credentials>@cluster0.tud3r.mongodb.net/
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({StudName: /^M/}).pretty();
[
  {
    _id: 1,
    StudName: 'MichelleJacintha',
    Grade: 'VII',
    Hobbies: 'InternetSurfing'
  }
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({StudName: /e/}).pretty();
[
  {
    _id: 1,
    StudName: 'MichelleJacintha',
    Grade: 'VII',
    Hobbies: 'InternetSurfing'
  }
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.count();
DeprecationWarning: Collection.count() is deprecated. Use countDocuments or estimatedDocumentCount.
2
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find().sort({StudName: -1}).pretty();
[
  {
    _id: 1,
    StudName: 'MichelleJacintha',
    Grade: 'VII',
    Hobbies: 'InternetSurfing'
  },
  {
    _id: 2, Grade: 'VII', StudName: 'AryanDavid', Hobbies: 'Chess' }
]
Atlas atlas-vvji9t-shard-0 [primary] myDB>
(To exit, press Ctrl+c again or Ctrl+d or type .exit)
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find().pretty();
[
  {
    _id: 1,
    StudName: 'MichelleJacintha',
    Grade: 'VII',
    Hobbies: 'InternetSurfing'
  },
  {
    _id: 2, Grade: 'VII', StudName: 'AryanDavid', Hobbies: 'Chess' }
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student3.find()
[
  {
    _id: ObjectId('67d0009c20453200f8c4ffcc'),
    StudName: 'AryanDavid',
    Grade: 'VIII',
    Hobbies: 'Chess'
  },
  {
    _id: ObjectId('67d0009c20453200f8c4ffd'),
    StudName: 'MichelleJacintha',
    Grade: 'VIII',
    Hobbies: 'Reading'
  }
]

Activities Terminal Mar 11 15:23
mongosh mongodb+srv://<credentials>@cluster0.tud3r.mongodb.net/
Hobbies: 'InternetSurfing'
}
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.save({StudName: "Vansl", Grade: "VI"})
TypeError: db.Student1.save is not a function
Atlas atlas-vvji9t-shard-0 [primary] myDB> use myDB;
already on db myDB
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.insertOne({ StudName: "Vansl", Grade: "VI" });
{
  acknowledged: true,
  insertedId: ObjectId('67d0031a8790ab5f31c00e7e')
}
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find()
[
  {
    _id: 1,
    StudName: 'MichelleJacintha',
    Grade: 'VII',
    Hobbies: 'InternetSurfing'
  },
  {
    _id: 2, Grade: 'VII', StudName: 'AryanDavid', Hobbies: 'Chess' },
  {
    _id: ObjectId('67d0031a8790ab5f31c00e7e'),
    StudName: 'Vansl',
    Grade: 'VI'
  }
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.insertOne({
  ...
  ...
  StudName: 'JohnDoe',
  ...
  Grade: 'VIII',
  ...
  Hobbies: 'Reading',
  ...
  Location: 'abc'
})
{
  acknowledged: true,
  insertedId: 4
}
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find()
[
  {
    _id: 1,
    StudName: 'MichelleJacintha',
    Grade: 'VII',
    Hobbies: 'InternetSurfing'
  },
  {
    _id: 2, Grade: 'VII', StudName: 'AryanDavid', Hobbies: 'Chess' },
  {
    _id: ObjectId('67d0031a8790ab5f31c00e7e'),
    StudName: 'Vansl',
    Grade: 'VI'
  },
  {
    _id: 4,
    StudName: 'JohnDoe',
    Grade: 'VIII',
    Hobbies: 'Reading',
    Location: 'abc'
  }
]

```

```

Activities Terminal Mar 11 15:23
mongosh mongodb+srv://<credentials>@cluster0.tud3r.mongodb.net/
[{"$set": {"Location": "abc"}}
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.updateOne(
... { "_id: 4 }, // Find the document with _id: 4
... { $set: { Location: "Network" } } // Set the Location field to "Network"
...
[
  acknowledged: true,
  insertedId: null,
  matchedCount: 1,
  modifiedCount: 1,
  upsertedCount: 0
}
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find()
[
  {
    "_id: 1,
    StudName: 'MichelleJacintha',
    Grade: 'VII',
    Hobbies: 'InternetSurfing'
  },
  { _id: 3, StudName: 'AryanDavid', Grade: 'VII', Hobbies: 'Chess' }
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find()
[
  {
    "_id: 1,
    StudName: 'MichelleJacintha',
    Grade: 'VII',
    Hobbies: 'InternetSurfing'
  },
  { _id: 3, Grade: 'VII', studName: 'AryanDavid', Hobbies: 'Chess' },
  {
    _id: ObjectId('67d0031a8790ab5f31c00e7e'),
    StudName: 'Vanst',
    Grade: 'VI'
  },
  {
    "_id: 4,
    StudName: 'JohnDoe',
    Grade: 'VIII',
    Hobbies: 'Reading',
    Location: 'Network'
  }
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({ _id: 1 }, { studName: 1, Grade: 1, _id: 0 });
[ { StudName: 'MichelleJacintha', Grade: 'VII' } ]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({Grade: { $eq: 'VII'}});
[
  {
    _id: ObjectId('67d0031a8790ab5f31c00e7e'),
    StudName: 'Vanst',
    Grade: 'VI'
  }
]

```

```

Activities Terminal Mar 11 15:24
mongosh mongodb+srv://<credentials>@cluster0.tud3r.mongodb.net/
[{"$set": {"Location": "abc"}}
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({studName: /$$/});
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.updateOne(
... { _id: 3 }, // Find the document with _id: 3
... { $set: { location: "Network" } } // Set the Location field to "Network"
...
[
  acknowledged: true,
  insertedId: null,
  matchedCount: 1,
  modifiedCount: 1,
  upsertedCount: 0
}
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find()
[
  {
    "_id: 1,
    StudName: 'MichelleJacintha',
    Grade: 'VII',
    Hobbies: 'InternetSurfing'
  },
  { _id: 3,
    Grade: 'VII',
    StudName: 'AryanDavid',
    Hobbies: 'Chess',
    Location: 'Network'
  },
  {
    _id: ObjectId('67d0031a8790ab5f31c00e7e'),
    StudName: 'Vanst',
    Grade: 'VI'
  },
  {
    "_id: 4,
    StudName: 'JohnDoe',
    Grade: 'VIII',
    Hobbies: 'Reading',
    Location: 'Network'
  }
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.updateOne({ _id: 3 }, { $set: { Location: null } });
{
  acknowledged: true,
  insertedId: null,
  matchedCount: 1,
  modifiedCount: 1,
  upsertedCount: 0
}

```

```

Activities Terminal Mar 11 15:24
mongosh mongodb+srv://<credentials>@cluster0.tud3r.mongodb.net/
[1]
-----+
] 0
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find()
[
  {
    "_id": 1,
    "StudName": "MichelleJacintha",
    "Grade": "VII",
    "Hobbies": "Internetsurfing"
  },
  {
    "_id": 2,
    "Grade": "VII",
    "StudName": "AryanDavid",
    "Hobbies": "Chess",
    "Location": null
  },
  {
    "_id": ObjectId("67d0031a8790ab5f31c00e7e"),
    "StudName": "Vanshi",
    "Grade": "VI"
  },
  {
    "_id": 4,
    "StudName": "JohnDoe",
    "Grade": "VIII",
    "Hobbies": "Reading",
    "Location": "Network"
  }
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.count();
4
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.countDocuments({Grade: "VII"});
2
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find({Grade: "VII"}).limit(3);
[
  {
    "_id": 1,
    "StudName": "MichelleJacintha",
    "Grade": "VII",
    "Hobbies": "Internetsurfing"
  },
  {
    "_id": 2,
    "Grade": "VII",
    "StudName": "AryanDavid",
    "Hobbies": "Chess",
    "Location": null
  }
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find().sort({StudName: 1});
[
  {
    "_id": 2,
    "Grade": "VII",
    "StudName": "AryanDavid"
  }
]

Activities Terminal Mar 11 15:24
mongosh mongodb+srv://<credentials>@cluster0.tud3r.mongodb.net/
[1]
-----+
] 0
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find().sort({StudName: 1});
[
  {
    "_id": 2,
    "Grade": "VII",
    "StudName": "AryanDavid",
    "Hobbies": "Chess",
    "Location": null
  },
  {
    "_id": 4,
    "StudName": "JohnDoe",
    "Grade": "VIII",
    "Hobbies": "Reading",
    "Location": "Network"
  },
  {
    "_id": 1,
    "StudName": "MichelleJacintha",
    "Grade": "VII",
    "Hobbies": "Internetsurfing"
  },
  {
    "_id": ObjectId("67d0031a8790ab5f31c00e7e"),
    "StudName": "Vanshi",
    "Grade": "VI"
  }
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.Student1.find().skip(2);
[
  {
    "_id": ObjectId("67d0031a8790ab5f31c00e7e"),
    "StudName": "Vanshi",
    "Grade": "VI"
  },
  {
    "_id": 4,
    "StudName": "JohnDoe",
    "Grade": "VIII",
    "Hobbies": "Reading",
    "Location": "Network"
  }
]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.food.insert({ _id: 1, fruits: ['grapes', 'mango', 'apple'] });
[ acknowledged: true, insertedId: '1' ]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.food.insert({ _id: 2, fruits: ['grapes', 'mango', 'cherry'] });
[ acknowledged: true, insertedId: '2' ]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.food.insert({ _id: 3, fruits: ['banana', 'mango'] });
[ acknowledged: true, insertedId: '3' ]
Atlas atlas-vvji9t-shard-0 [primary] myDB>
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.food.find()
[
  { _id: 1, fruits: ['grapes', 'mango', 'apple'] },
  { _id: 2, fruits: ['grapes', 'mango', 'cherry'] }
]

```

```

Activities Terminal Mar 11 15:24
mongosh mongodb+srv://<credentials>@cluster0.tud3r.mongodb.net/
[{"_id": 2, "fruits": ["grapes", "mango", "cherry"]},
 {"_id": 3, "fruits": ["banana", "mango"]}]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.food.find({ fruits: ['grapes', 'mango', 'apple'] });
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.food.find({ "fruits": 1: "grapes" });

Atlas atlas-vvji9t-shard-0 [primary] myDB> db.food.find({ "fruits": {$size: 2} });
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.food.find({ "_id": 1 }, {"fruits": {$slice: 2}});
[{"_id": 1, "fruits": ["grapes", "mango"]}]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.food.find({ "fruits": {$all: ["mango", "grapes"]} });
[{"_id": 1, "fruits": ["grapes", "mango", "apple"]},
 {"_id": 2, "fruits": ["grapes", "cherry"]}]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.food.find({ "fruits.0": "grapes" });
[{"_id": 1, "fruits": ["grapes", "mango", "apple"]},
 {"_id": 2, "fruits": ["grapes", "mango", "cherry"]}]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.food.find();
[{"_id": 1, "fruits": ["grapes", "mango", "apple"]},
 {"_id": 2, "fruits": ["grapes", "mango", "cherry"]},
 {"_id": 3, "fruits": ["banana", "mango"]}]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.food.update({_id: 3}, {"$set: {"fruits.1": "apple"}});
{
  acknowledged: true,
  insertedId: null,
  matchedCount: 1,
  modifiedCount: 1,
  upsertedCount: 0
}
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.food.find();
[{"_id": 1, "fruits": ["grapes", "mango", "apple"]},
 {"_id": 2, "fruits": ["grapes", "mango", "cherry"]},
 {"_id": 3, "fruits": ["banana", "apple"]}]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.food.update({_id: 2}, {"$push: {price: {grapes: 80, mango: 200, cherry: 100}}"});
{
  acknowledged: true,
  insertedId: null,
  matchedCount: 1,
  modifiedCount: 1,
  upsertedCount: 0
}
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.food.find();
[{"_id": 1, "fruits": ["grapes", "mango", "apple"]},
 {"_id": 2, "fruits": ["grapes", "mango", "cherry"],
   price: [ {grapes: 80, mango: 200, cherry: 100}]},
 {"_id": 3, "fruits": ["banana", "apple"]}]
Activities Terminal Mar 11 15:25
mongosh mongodb+srv://<credentials>@cluster0.tud3r.mongodb.net/
[{"price: [{grapes: 80, mango: 200, cherry: 100}]},
 {"_id: 3, "fruits": ["banana", "apple"]}]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.food.update({_id: 1}, {"$push: {"fruits": 1}});
{
  acknowledged: true,
  insertedId: null,
  matchedCount: 1,
  modifiedCount: 1,
  upsertedCount: 0
}
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.food.find();
[{"_id": 1, "fruits": ["grapes", "mango"]},
 {"_id": 2, "fruits: [{grapes: 80, mango: 200, cherry: 100}]},
 {"_id": 3, "fruits": ["banana", "apple"]}]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.food.update({_id: 1}, {"$addToSet: {"fruits": "orange"}});
{
  acknowledged: true,
  insertedId: null,
  matchedCount: 1,
  modifiedCount: 1,
  upsertedCount: 0
}
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.food.find();
[{"_id": 1, "fruits": ["grapes", "mango", "orange"]},
 {"_id": 2, "fruits: [{grapes: 80, mango: 200, cherry: 100}]},
 {"_id": 3, "fruits": ["banana", "apple"]}]
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.food.update({_id: 2}, {"$pullAll: {"fruits": ["grapes", "mango"]}});
{
  acknowledged: true,
  insertedId: null,
  matchedCount: 1,
  modifiedCount: 1,
  upsertedCount: 0
}
Atlas atlas-vvji9t-shard-0 [primary] myDB> db.food.find();
[{"_id": 1, "fruits": ["grapes", "mango", "orange"]},
 {"_id": 2, "fruits: [{cherry: 100}]},
 {"_id": 3, "fruits": ["banana", "apple"]}]

```

```
Activities Terminal Mar 11 15:25
mongosh mongodb+srv://[credentials]@cluster0.tud3r.mongodb.net/
[{"price": [ { grapes: 80, mango: 200, cherry: 100 } ]}, {"_id: 3, fruits: [ 'banana', 'apple' ]}], {"acknowledged: true, insertedIds: null, matchedCount: 1, modifiedCount: 1, upsertedCount: 0}], {"_id: 1, fruits: [ 'grapes', 'mango', 'orange' ]}, {"_id: 2, fruits: [ 'cherry' ], price: [ { grapes: 80, mango: 200, cherry: 100 } ]}, {"_id: 3, fruits: [ 'apple' ]}], {"ok: 1}], {"_id: 1, AcctBal: 1000, AcctType: "S"}, {"_id: 2, AcctBal: 1500, AcctType: "S"}, {"_id: 3, AcctBal: 1200, AcctType: "C"}, {"_id: 4, AcctBal: 800, AcctType: "S"}, {"_id: 5, AcctBal: 500, AcctType: "S"}], [{"acknowledged: true, insertedIds: [ '_id: 1, ObjectId('67d007098790ab5f31c00e7ff'), '_id: 2, ObjectId('67d007098790ab5f31c00e80'), '_id: 3, ObjectId('67d007098790ab5f31c00e81'), '_id: 4, ObjectId('67d007098790ab5f31c00e82'), '_id: 5, ObjectId('67d007098790ab5f31c00e83') }]}], [{"_id: ObjectId('67d007098790ab5f31c00e7ff'), custId: 1, AcctBal: 1000, AcctType: "S"}, {"_id: ObjectId('67d007098790ab5f31c00e80'), custId: 2, AcctBal: 1500, AcctType: "S"}, {"_id: ObjectId('67d007098790ab5f31c00e81'), custId: 3, AcctBal: 1200, AcctType: "C"}, {"_id: ObjectId('67d007098790ab5f31c00e82'), custId: 4, AcctBal: 800, AcctType: "S"}, {"_id: ObjectId('67d007098790ab5f31c00e83'), custId: 5, AcctBal: 500, AcctType: "S"}]]
```

```
Activities ┌───────── Terminal ──────────┐ Mar 11 15:25
mongosh mongodb+srv://<credentials>@cluster0.tud3r.mongodb.net/
Atlas atlas-vvji9l-shard-0 [primary] myDB> db.Customers.find()
[{"_id": ObjectId("67d007098790ab5f31c00e7f"), "custID": 1, "AcctBal": 1000, "AcctType": "S"}, {"_id": ObjectId("67d007098790ab5f31c00e80"), "custID": 2, "AcctBal": 1500, "AcctType": "S"}, {"_id": ObjectId("67d007098790ab5f31c00e81"), "custID": 1, "AcctBal": 1200, "AcctType": "C"}, {"_id": ObjectId("67d007098790ab5f31c00e82"), "custID": 3, "AcctBal": 800, "AcctType": "S"}, {"_id": ObjectId("67d007098790ab5f31c00e83"), "custID": 2, "AcctBal": 500, "AcctType": "S"}]
Atlas atlas-vvji9l-shard-0 [primary] myDB> db.Customers.aggregate([
... {
...   $group: {
...     id: "$custID",
...     TotAccBal: { $sum: "$AcctBal" }
...   }
... }
... ]);
[{"_id": 2, "TotAccBal": 2000}, {"_id": 3, "TotAccBal": 800}, {"_id": 1, "TotAccBal": 2200}]
Atlas atlas-vvji9l-shard-0 [primary] myDB> db.Customers.aggregate([
... {
...   $group: {
...     id: "$custID",
...     TotAccBal: { $sum: "$AcctBal" }
...   }
... }
... ]);

```

```

Activities Terminal Mar 11 15:26
bmseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: ~
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoexport
2025-03-11T14:39:05.282+0530 must specify a collection
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ try "mongoexport --help" for more information
mongosh command not found
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ ~mongosh "mongodb+srv://cluster0.tud3r.mongodb.net/" --apiVersion 1 --username mariacs22 --password 8Ja9u0zIrEjzBlot
-mongosh: command not found
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ mongosh "mongodb+srv://cluster0.tud3r.mongodb.net/" --apiVersion 1 --username mariacs22 --password 8Ja9u0zIrEjzBlot
Current Mongosh Log ID: 67cffdfcfa3c80154c00e7d
Connecting to: mongodb+srv://credentials@cluster0.tud3r.mongodb.net/?appName=mongosh+2.2.1
Using MongoDB: 8.0.5 (API Version 1)
Using Mongosh: 2.2.1
mongosh 2.4.2 is available for download: https://www.mongodb.com/try/download/shell
For mongosh info see: https://docs.mongodb.com/mongosh-shell/
Atlas atlas-vvji9l-shard-0 [primary] test> use myDB;
switched to db myDB
Atlas atlas-vvji9l-shard-0 [primary] myDB> mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "StudentName,Grade,Hobbies"
Unccaught:
SyntaxError: Missing semicolon. (1:14)
> 1 | mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "StudentName,Grade,Hobbies"
2 |
Atlas atlas-vvji9l-shard-0 [primary] myDB> mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "StudentName,Grade,Hobbies"
Unccaught:
SyntaxError: Missing semicolon. (1:14)
> 1 | mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "StudentName,Grade,Hobbies"
2 |
Atlas atlas-vvji9l-shard-0 [primary] myDB> (To exit, press Ctrl+c again or Ctrl+d or type .exit)
Atlas atlas-vvji9l-shard-0 [primary] myDB> mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "StudentName,Grade,Hobbies"
Unccaught:
SyntaxError: Missing semicolon. (1:14)
> 1 | mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "StudentName,Grade,Hobbies"
2 |
Atlas atlas-vvji9l-shard-0 [primary] myDB> exit
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "StudentName,Grade,Hobbies"
2025-03-11T14:42:20.815+0530 connected to: mongodb+srv://[*REDACTED*]@cluster0.tud3r.mongodb.net/myDB
2025-03-11T14:42:21.068+0530 exported 2 records
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ 
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "StudentName,Grade,Hobbies"
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: ~

Activities Terminal Mar 11 15:26
bmseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: ~
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ 
Atlas atlas-vvji9l-shard-0 [primary] myDB> (To exit, press Ctrl+c again or Ctrl+d or type .exit)
Atlas atlas-vvji9l-shard-0 [primary] myDB> mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "StudentName,Grade,Hobbies"
Unccaught:
SyntaxError: Missing semicolon. (1:14)
> 1 | mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "StudentName,Grade,Hobbies"
2 |
Atlas atlas-vvji9l-shard-0 [primary] myDB> exit
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "StudentName,Grade,Hobbies"
2025-03-11T14:42:20.815+0530 connected to: mongodb+srv://[*REDACTED*]@cluster0.tud3r.mongodb.net/myDB
2025-03-11T14:42:21.068+0530 exported 2 records
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ 
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "StudentName,Grade,Hobbies"
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: ~

? 2025-03-11T14:43:35.174+0530 connected to: mongodb+srv://[*REDACTED*]@cluster0.tud3r.mongodb.net/myDB
2025-03-11T14:43:35.437+0530 exported 2 records
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ 
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ 
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ 
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "StudentName,Grade,Hobbies"
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: ~

? 2025-03-11T14:44:44.222+0530 connected to: mongodb+srv://[*REDACTED*]@cluster0.tud3r.mongodb.net/myDB
2025-03-11T14:44:44.529+0530 exported 2 records
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ whoami
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: ~
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "/home/bmseccse/Desktop/qwer.csv" --fields "StudentName,Grade,Hobbies"
2025-03-11T14:47:52.993+0530 connected to: mongodb+srv://[*REDACTED*]@cluster0.tud3r.mongodb.net/myDB
2025-03-11T14:47:53.005+0530 exported 2 records
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoimport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student3 --type csv --file "/home/bmseccse/Desktop/qwer.csv" --headerline
2025-03-11T14:51:32.165+0530 connected to: mongodb+srv://[*REDACTED*]@cluster0.tud3r.mongodb.net/myDB
2025-03-11T14:51:32.316+0530 2 document(s) imported successfully. 0 document(s) failed to import.
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ 
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoimport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --headerline --file "/Downloads/airline.csv"
2025-03-11T14:55:09.112+0530 Failed: open /home/bmseccse/Downloads/airline.csv: no such file or directory
2025-03-11T14:55:09.112+0530 0 document(s) imported successfully. 0 document(s) failed to import.
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoimport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --headerline --file "/Downloads/airline.csv"
2025-03-11T14:56:05.413+0530 Failed: open /home/bmseccse/Desktop/airline.csv: no such file or directory
2025-03-11T14:56:05.413+0530 0 document(s) imported successfully. 0 document(s) failed to import.
bnseccse@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoexport --uri "mongodb+srv://mariacs22:8Ja9u0zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "/Downloads/airline.csv" --fields "Year" "Quarter"

```

```
Activities Terminal Mar 11 15:26
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC ~

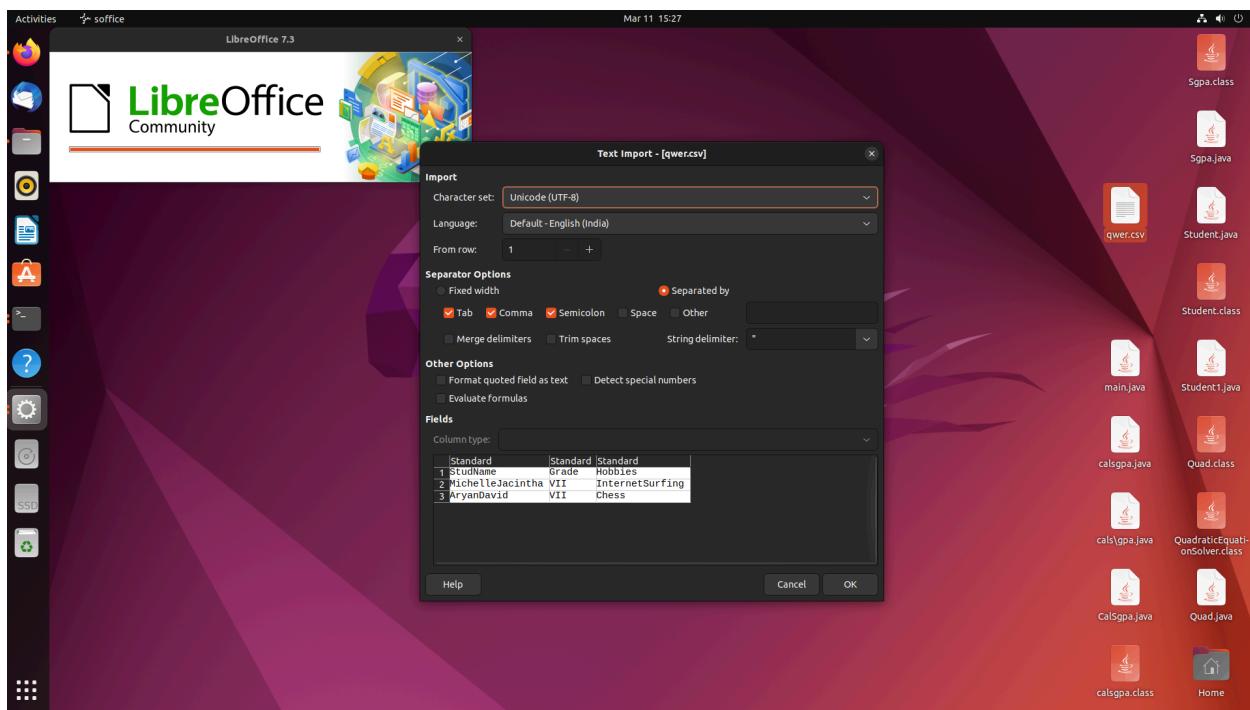
dname,Grade,Hobbies
Uncought:
SyntaxError: Missing semicolon. (1:14)
> 1 | mongoreport --url "mongodb+srv://marlacs22:8J3a9u6zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "StudName,Grade,Hobbies"
2 |

Atlas atlas-vv191-shard-0 [primary] myDB> exit
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoexport -url "mongodb+srv://marlacs22:8J3a9u6zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/output.txt" --fields "StudName,Grade,Hobbies"

2025-03-11T14:45:20.815+0530 connected to: mongodb+srv://[*REDACTED*]@cluster0.tud3r.mongodb.net/myDB
2025-03-11T14:45:21.868+0530 exported 2 records
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: $ 
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoexport -url "mongodb+srv://marlacs22:8J3a9u6zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/out.txt" --file "StudName,Grade,Hobbies"
ld5 "StudName,Grade,Hobbies"

2025-03-11T14:43:35.174+0530 connected to: mongodb+srv://[*REDACTED*]@cluster0.tud3r.mongodb.net/myDB
2025-03-11T14:43:35.437+0530 exported 2 records
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: $ 
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: $ 
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoexport -url "mongodb+srv://marlacs22:8J3a9u6zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "D:/qwer.txt" --fields "StudName,Grade,Hobbies"
2025-03-11T14:44:44.22+0530 connected to: mongodb+srv://[*REDACTED*]@cluster0.tud3r.mongodb.net/myDB
2025-03-11T14:44:44.529+0530 exported 2 records
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: $ whoami
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: $ ^C
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoexport -url "mongodb+srv://marlacs22:8J3a9u6zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --out "/home/bmscsece/Desktop/qwer.csv" --fields "StudName,Grade,Hobbies"
2025-03-11T14:47:52.993+0530 connected to: mongodb+srv://[*REDACTED*]@cluster0.tud3r.mongodb.net/myDB
2025-03-11T14:47:53.277+0530 exported 2 records
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoimport --url "mongodb+srv://marlacs22:8J3a9u6zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student3 --type csv --file "/home/bmscsece/Desktop/qwer.csv" --headerline
2025-03-11T14:51:32.165+0530 connected to: mongodb+srv://[*REDACTED*]@cluster0.tud3r.mongodb.net/myDB

2025-03-11T14:51:32.316+0530 2 document(s) imported successfully. 0 document(s) failed to import.
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: $ 
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoimport --url "mongodb+srv://marlacs22:8J3a9u6zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --headerline --file "/Downloads/airline.csv"
2025-03-11T14:55:09.112+0530 Failed: open /home/bmscsece/Downloads/airline.csv: no such file or directory
2025-03-11T14:55:09.112+0530 0 document(s) imported successfully. 0 document(s) failed to import.
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoimport --url "mongodb+srv://marlacs22:8J3a9u6zIrEjzBlot@cluster0.tud3r.mongodb.net/myDB" --collection Student1 --type csv --headerline --file "/Desktop/airline.csv"
2025-03-11T14:55:09.413+0530 Failed: open /home/bmscsece/Desktop/airline.csv: no such file or directory
2025-03-11T14:55:09.413+0530 0 document(s) imported successfully. 0 document(s) failed to import.
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: $ mongoexport --host localhost --db myDB --collection Student1 --type=csv --out ~/Downloads/airline.csv --fields "Year","Quarter"
2025-03-11T14:58:48.870+0530 connected to: mongodb+srv://localhost/
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: $ exported 0 records
```



New_S... +

cluster0.tud3r.mongodb.net > test > New_Student

Documents 6 Aggregations Schema Indexes 1 Validation

Type a query: { field: 'value' } or [Generate query](#)

Add DATA EXPORT DATA UPDATE DELETE Explain Reset Find Options

26 1–6 of 6

`_id: ObjectId('67c6c0281bcf8a6e28fa4210')`
RollNo : 3
Age : 21
Cont : 5576
email : "anubhav.de9@gmail.com"

`_id: ObjectId('67c6c01e1bcf8a6e28fa4215')`
RollNo : 2
Age : 22
Cont : 9976
email : "anushka.de9@gmail.com"

`_id: ObjectId('67c6c0141bcf8a6e28fa4214')`
RollNo : 1
Age : 21
Cont : 9876
email : "antara.de9@gmail.com"

`_id: ObjectId('67c6c05d1bcf8a6e28fa4219')`
RollNo : 11
Age : 22
Name : "FEM"
Cont : 2276
email : "rea.de9@gmail.com"

`_id: ObjectId('67c6c02f1bcf8a6e28fa4217')`
RollNo : 4
Age : 20
Cont : 4476
email : "pani.de9@gmail.com"

`_id: ObjectId('67c6c0351bcf8a6e28fa4218')`
RollNo : 10
Age : 20
Cont : 2276
email : "Abhinav@gmail.com"

`_id: ObjectId('67c6c0141bcf8a6e28fa4215')`
RollNo : 3
Age : 21
Cont : 5576
email : "anubhav.de9@gmail.com"

`_id: ObjectId('67c6c0141bcf8a6e28fa4214')`
RollNo : 1
Age : 22
Cont : 9876
email : "antara.de9@gmail.com"

`_id: ObjectId('67c6c05d1bcf8a6e28fa4219')`
RollNo : 11
Age : 22
Name : "FEM"
Cont : 2276
email : "rea.de9@gmail.com"

`_id: ObjectId('67c6c02f1bcf8a6e28fa4217')`
RollNo : 4
Age : 20
Cont : 4476
email : "pani.de9@gmail.com"

`_id: ObjectId('67c6c0351bcf8a6e28fa4218')`
RollNo : 10
Age : 23
Cont : 2276
email : "Abhinav@gmail.com"

ENG IN 15:20:44 04-03-2025

abc - Excel

File Home Insert Page Layout Formulas Data Review View Help eScan DLP Classification

POSSIBLE DATA LOSS: Some features might be lost if you save this workbook in the comma-delimited (.csv) format. To preserve these features, save it in an Excel file format.

Don't show again Save As...

A1

```
[{"_id": {"$oid": "5c0d0141bc8a6e28fa4214"}, "RollNo": "1", "Age": "21", "Cont": "9876", "email": "anurara.de@gmail.com"}, {"_id": {"$oid": "5c0d0141bc8a6e28fa4215"}, "RollNo": "2", "Age": "22", "Cont": "9976", "email": "anushka.de@gmail.com"}, {"_id": {"$oid": "5c0d0141bc8a6e28fa4216"}, "RollNo": "3", "Age": "21", "Cont": "5576", "email": "anubhav.de@gmail.com"}, {"_id": {"$oid": "5c0d0141bc8a6e28fa4217"}, "RollNo": "4", "Age": "20", "Cont": "4476", "email": "pani.de@gmail.com"}, {"_id": {"$oid": "5c0d0141bc8a6e28fa4218"}, "RollNo": "10", "Age": "23", "Cont": "2276", "email": "Abhinav@gmail.com"}, {"_id": {"$oid": "5c0d0141bc8a6e28fa4219"}, "RollNo": "11", "Age": "22", "Name": "FEM", "Cont": "2276", "email": "rea.de@gmail.com"}]
```

Ready 32°C Partly sunny

File Edit Selection View Go Run Terminal Help

Release Notes: 1.97.2

```
C:\Users> STUDENT > ejfjson > _
```

```
[{"_id": {"$oid": "67c6c0351bcf8a6e28fa4214"}, "RollNo": "1", "Age": "21", "Cont": "9876", "email": "anurara.de@gmail.com"}, {"_id": {"$oid": "67c6c0351bcf8a6e28fa4215"}, "RollNo": "2", "Age": "22", "Cont": "9976", "email": "anushka.de@gmail.com"}, {"_id": {"$oid": "67c6c0351bcf8a6e28fa4216"}, "RollNo": "3", "Age": "21", "Cont": "5576", "email": "anubhav.de@gmail.com"}, {"_id": {"$oid": "67c6c0351bcf8a6e28fa4217"}, "RollNo": "4", "Age": "20", "Cont": "4476", "email": "pani.de@gmail.com"}, {"_id": {"$oid": "67c6c0351bcf8a6e28fa4218"}, "RollNo": "10", "Age": "23", "Cont": "2276", "email": "Abhinav@gmail.com"}, {"_id": {"$oid": "67c6c0351bcf8a6e28fa4219"}, "RollNo": "11", "Age": "22", "Name": "FEM", "Cont": "2276", "email": "rea.de@gmail.com"}]
```

Log in to Sourcery

Source: Sourcery Log in via web browser Log in via token

Please sign in to start using Tabnine

Source: Tabnine AI Sign in

Launched 1 △ 0 Sourcery

Ln 1, Col 1 Spaces: 4 UTF-8 LF JSON Go Live Tabnine: Sign-in is required

32°C Partly sunny

File Home Insert Page Layout Formulas Data Review View Help eScan DLP Classification

POSSIBLE DATA LOSS: Some features might be lost if you save this workbook in the comma-delimited (.csv) format. To preserve these features, save it in an Excel file format.

Don't show again Save As...

A1

```
[{"_id": {"$oid": "5c0d0141bc8a6e28fa4214"}, "RollNo": "1", "Age": "21", "Cont": "9876", "email": "anurara.de@gmail.com"}, {"_id": {"$oid": "5c0d0141bc8a6e28fa4215"}, "RollNo": "2", "Age": "22", "Cont": "9976", "email": "anushka.de@gmail.com"}, {"_id": {"$oid": "5c0d0141bc8a6e28fa4216"}, "RollNo": "3", "Age": "21", "Cont": "5576", "email": "anubhav.de@gmail.com"}, {"_id": {"$oid": "5c0d0141bc8a6e28fa4217"}, "RollNo": "4", "Age": "20", "Cont": "4476", "email": "pani.de@gmail.com"}, {"_id": {"$oid": "5c0d0141bc8a6e28fa4218"}, "RollNo": "10", "Age": "23", "Cont": "2276", "email": "Abhinav@gmail.com"}, {"_id": {"$oid": "5c0d0141bc8a6e28fa4219"}, "RollNo": "11", "Age": "22", "Name": "FEM", "Cont": "2276", "email": "rea.de@gmail.com"}]
```

Ready 32°C Partly sunny

File Edit Selection View Go Run Terminal Help

Release Notes: 1.97.2

```
C:\Users> STUDENT > ejfjson > _
```

```
[{"_id": {"$oid": "67c6c0351bcf8a6e28fa4214"}, "RollNo": "1", "Age": "21", "Cont": "9876", "email": "anurara.de@gmail.com"}, {"_id": {"$oid": "67c6c0351bcf8a6e28fa4215"}, "RollNo": "2", "Age": "22", "Cont": "9976", "email": "anushka.de@gmail.com"}, {"_id": {"$oid": "67c6c0351bcf8a6e28fa4216"}, "RollNo": "3", "Age": "21", "Cont": "5576", "email": "anubhav.de@gmail.com"}, {"_id": {"$oid": "67c6c0351bcf8a6e28fa4217"}, "RollNo": "4", "Age": "20", "Cont": "4476", "email": "pani.de@gmail.com"}, {"_id": {"$oid": "67c6c0351bcf8a6e28fa4218"}, "RollNo": "10", "Age": "23", "Cont": "2276", "email": "Abhinav@gmail.com"}, {"_id": {"$oid": "67c6c0351bcf8a6e28fa4219"}, "RollNo": "11", "Age": "22", "Name": "FEM", "Cont": "2276", "email": "rea.de@gmail.com"}]
```

Log in to Sourcery

Source: Sourcery Log in via web browser Log in via token

Please sign in to start using Tabnine

Source: Tabnine AI Sign in

Launched 1 △ 0 Sourcery

Ln 1, Col 1 Spaces: 4 UTF-8 LF JSON Go Live Tabnine: Sign-in is required

32°C Partly sunny

```
Command Prompt × + ▾ - □ ×
Microsoft Windows [Version 10.0.22631.4890]
(c) Microsoft Corporation. All rights reserved.

C:\Users\STUDENT>mongoexport --uri="mongodb+srv://mariacs22:8Ja9u6zIrEjzBlot@cluster0.tud3r.mongodb.net/test" --collection=Student --out=abc.csv
2025-03-04T15:09:05.137+0530      connected to: mongodb+srv://[**REDACTED**]@cluster0.tud3r.mongodb.net/test
2025-03-04T15:09:05.359+0530      exported 6 records

C:\Users\STUDENT>mongoexport --uri="mongodb+srv://mariacs22:8Ja9u6zIrEjzBlot@cluster0.tud3r.mongodb.net/test" --collection=Student --out=efg.json
2025-03-04T15:14:50.067+0530      connected to: mongodb+srv://[**REDACTED**]@cluster0.tud3r.mongodb.net/test
2025-03-04T15:14:50.288+0530      exported 6 records

C:\Users\STUDENT>mongoimport --uri="mongodb+srv://mariacs22:8Ja9u6zIrEjzBlot@cluster0.tud3r.mongodb.net/test" --collection>New_Student --type=json --file="C:\Users\STUDENT\efg.json"
2025-03-04T15:15:35.501+0530      connected to: mongodb+srv://[**REDACTED**]@cluster0.tud3r.mongodb.net/test
2025-03-04T15:15:35.590+0530      6 document(s) imported successfully. 0 document(s) failed to import.

C:\Users\STUDENT>
```

EXPERIMENT-2

Neo4j

Code with Screenshots:

The screenshot shows the Neo4j browser interface with two main sections: a code editor on the left and a results panel on the right.

Code Editor (Left):

- Panel title: Table
- Code input field:

```
neo4j$ CREATE (a:STUDENT {name:"Ram" , id: 12})
```
- Output message: Added 1 label, created 1 node, set 2 properties, completed after 27 ms.

Results Panel (Right):

- Panel title: Overview
- Node labels:
 - * (1) STUDENT (1)
- Message: Displaying 1 nodes, 0 relationships.
- Visual representation: A single pink circular node labeled "Ram".
- Control icons: A vertical toolbar on the right with icons for search, refresh, and other operations.

```
neo4j$ CREATE (a:STUDENT {name:"Abhi" , id: 14})
```

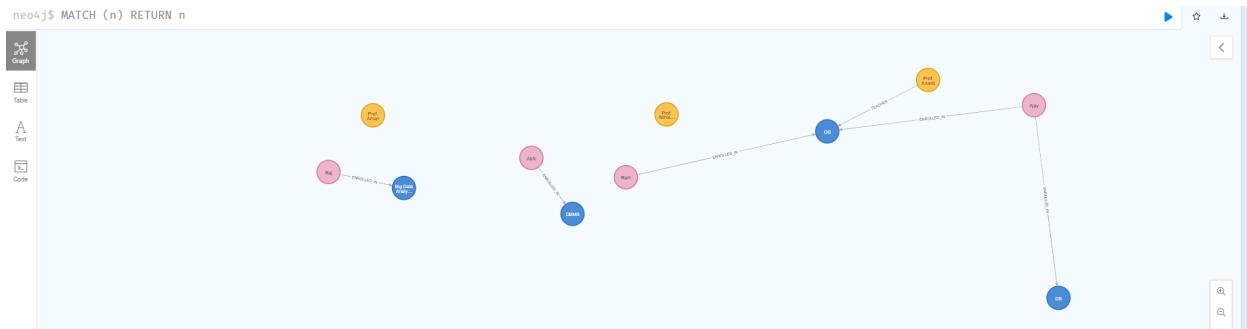
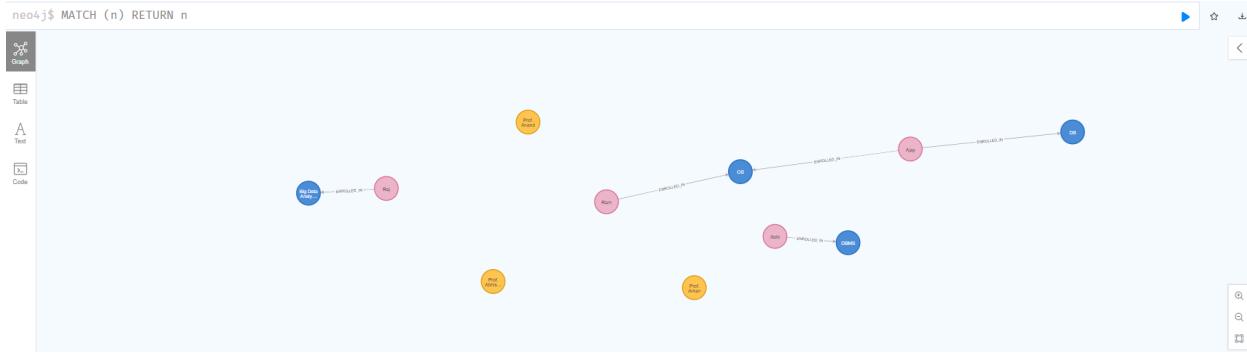
Added 1 label, created 1 node, set 2 properties, completed after 6 ms.

```
neo4j$ CREATE (a:STUDENT {name:"Raj" , id: 13})
```

Added 1 label, created 1 node, set 2 properties, completed after 6 ms.

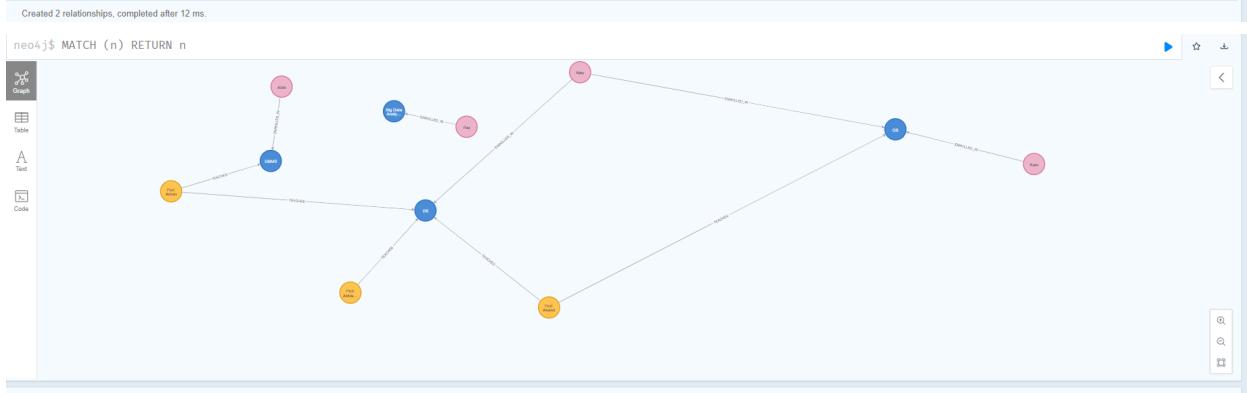
```
neo4j$ CREATE (a:STUDENT {name:"Ajay" , id: 15})
```

Added 1 label, created 1 node, set 2 properties, completed after 6 ms.



neo4j\$ MATCH (a:TEACHER) , (b:COURSE) where a.name = "Prof. Anand" AND b.name = "OS" CREATE (a)-[r:TEACHES]-(b)

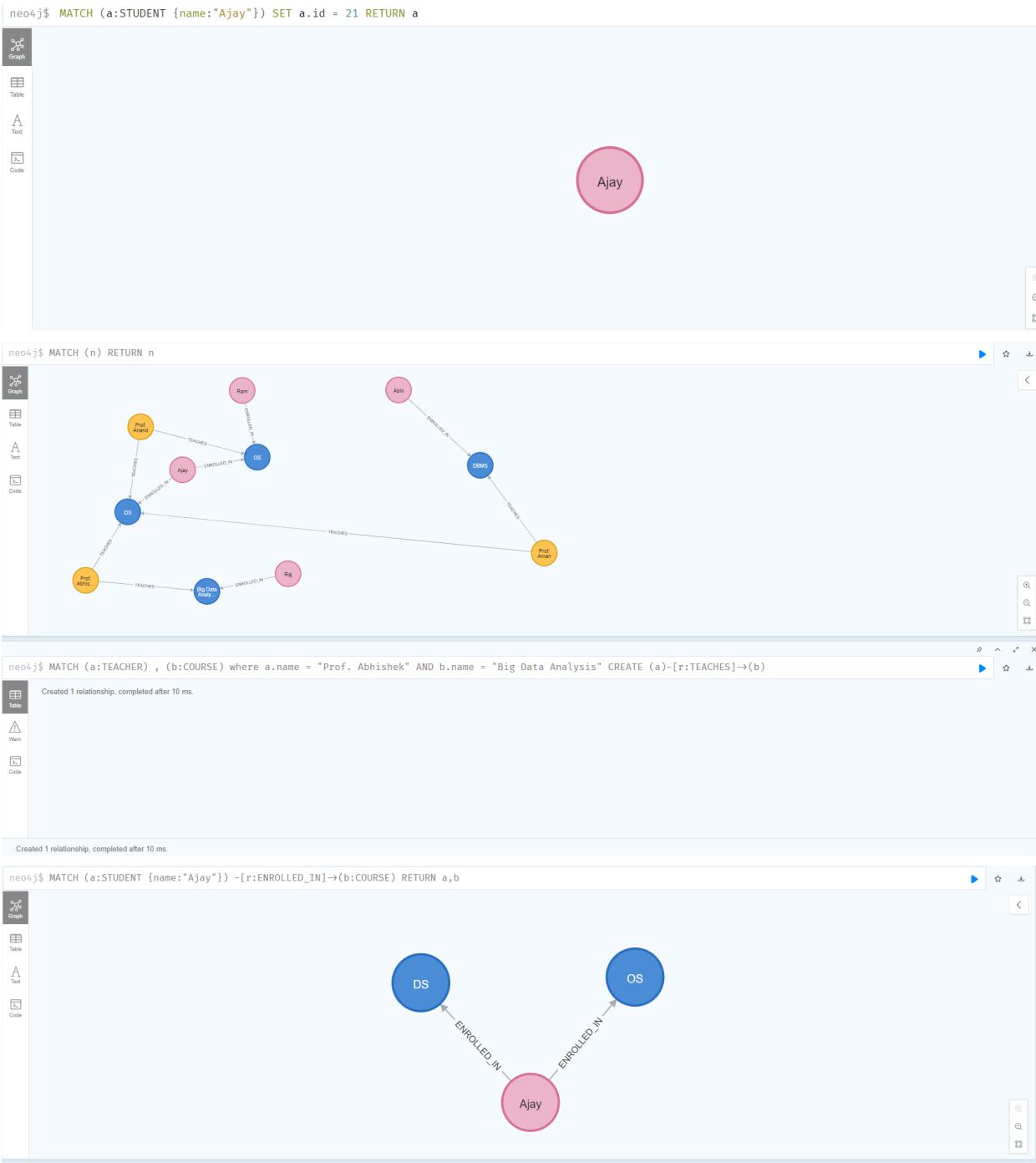
Created 2 relationships, completed after 12 ms.

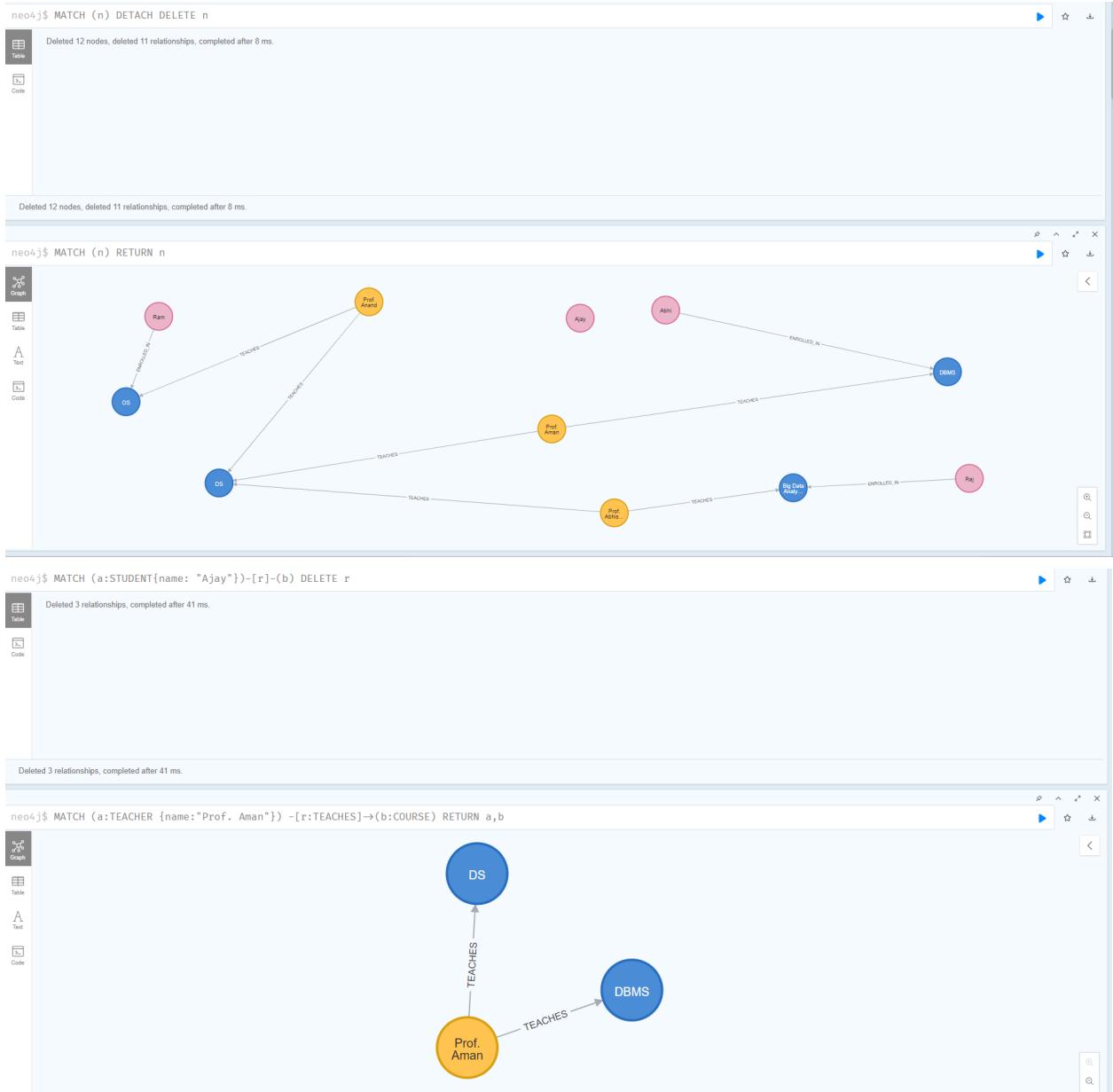


neo4j\$ MATCH (a:TEACHER) , (b:COURSE) where a.name = "Prof. Aman" AND b.name = "DBMS" or b.name = "DS" CREATE (a)-[r:TEACHES]-(b)

Created 4 relationships, completed after 12 ms.

Created 4 relationships, completed after 12 ms





EXPERIMENT-3

Perform the following DB operations using Cassandra.

- Create a keyspace by name Employee
- Create a column family by name Employee-Info with attributes
Emp_Id Primary Key, Emp_Name,
Designation, Date_of_Joining,
Salary, Dept_Name
- Insert the values into the table in batch
- Update Employee name and Department of Emp-Id 121
- Sort the details of Employee records based on salary
- Alter the schema of the table Employee_Info to add a column Projects which stores a set of Projects done by the corresponding Employee.
- Update the altered table to add project names.
- Create a TTL of 15 seconds to display the values of Employees.

CODE WITH SCREENSHOTS:

```
Activities Terminal Apr 8 14:12
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC:~ cqlsh
[travis@bmscsece ~]$ cqlsh 127.0.0.1:9042
[cqlsh 6.1.0 | Cassandra 4.1.6 | CQL spec 3.4.6 | Native protocol v5]
Use HELP for help.
cqlsh> CREATE KEYSPACE Employee
...   WITH replication = {'class': 'SimpleStrategy', 'replication_factor': 1};

cqlsh>
cqlsh> USE Employee;
cqlsh> employee>
cqlsh:employee> CREATE TABLE Employee_Info (
...     Emp_Id int PRIMARY KEY,
...     Emp_Name text,
...     Designation text,
...     Date_of_Joining date,
...     Salary double,
...     Dept_Name text
... );

cqlsh:employees>
cqlsh:employees> BEGIN BATCH
...     INSERT INTO Employee_Info (Emp_Id, Emp_Name, Designation, Date_of_Joining, Salary, Dept_Name)
...     VALUES (121, 'Alice', 'Manager', '2020-05-01', 75000, 'HR');
...     INSERT INTO Employee_Info (Emp_Id, Emp_Name, Designation, Date_of_Joining, Salary, Dept_Name)
...     VALUES (122, 'Bob', 'Developer', '2021-01-15', 60000, 'IT');
...     INSERT INTO Employee_Info (Emp_Id, Emp_Name, Designation, Date_of_Joining, Salary, Dept_Name)
...     VALUES (123, 'Charlie', 'Analyst', '2019-09-10', 50000, 'Finance');
...     APPLY BATCH;
cqlsh:employees>
cqlsh:employees> SELECT * FROM Employee_Info;

emp_id | date_of_joining | dept_name | designation | emp_name | salary
-----+-----------------+-----------+-------------+----------+-----
123 | 2019-09-10 | Finance | Analyst | Charlie | 50000
122 | 2021-01-15 | IT | Developer | Bob | 60000
121 | 2020-05-01 | HR | Manager | Alice | 75000
(3 rows)

cqlsh:employees> UPDATE Employee_Info SET Emp_Name = 'Alicia', Dept_Name = 'Admin' WHERE Emp_Id = 121;
cqlsh:employees> SELECT * FROM Employee_Info;

emp_id | date_of_joining | dept_name | designation | emp_name | salary
-----+-----------------+-----------+-------------+----------+-----
123 | 2019-09-10 | Finance | Analyst | Charlie | 50000
122 | 2021-01-15 | IT | Developer | Bob | 60000
121 | 2020-05-01 | Admin | Manager | Alicia | 75000
(3 rows)

cqlsh:employees> SELECT * FROM Employee_Info WHERE Salary >= 0 ALLOW FILTERING;

emp_id | date_of_joining | dept_name | designation | emp_name | salary
-----+-----------------+-----------+-------------+----------+-----
123 | 2019-09-10 | Finance | Analyst | Charlie | 50000
122 | 2021-01-15 | IT | Developer | Bob | 60000
121 | 2020-05-01 | Admin | Manager | Alicia | 75000
```

```

Activities Terminal Apr 8 14:12
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~
121 | 2020-05-01 | HR | Manager | Alice | 75000
(3 rows)
cqsh:employee> UPDATE Employee_Info SET Emp_Name = 'Alicia', Dept_Name = 'Admin' WHERE Emp_Id = 121;
cqsh:employee> SELECT*FROM Employee_Info;
emp_id | date_of_joining | dept_name | designation | emp_name | salary
-----+-----+-----+-----+-----+
123 | 2019-09-10 | Finance | Analyst | Charlie | 50000
122 | 2021-01-15 | IT | Developer | Bob | 60000
121 | 2020-05-01 | Admin | Manager | Alicia | 75000
(3 rows)
cqsh:employee> SELECT * FROM Employee_Info WHERE Salary >= 0 ALLOW FILTERING;
emp_id | date_of_joining | dept_name | designation | emp_name | salary
-----+-----+-----+-----+-----+
123 | 2019-09-10 | Finance | Analyst | Charlie | 50000
122 | 2021-01-15 | IT | Developer | Bob | 60000
121 | 2020-05-01 | Admin | Manager | Alicia | 75000
(3 rows)
cqsh:employee> ALTER TABLE Employee_Info ADD Projects set<text>;
cqsh:employee> SELECT*FROM Employee_Info;
SSD emp_id | date_of_joining | dept_name | designation | emp_name | projects | salary
-----+-----+-----+-----+-----+-----+
123 | 2019-09-10 | Finance | Analyst | Charlie | null | 50000
122 | 2021-01-15 | IT | Developer | Bob | null | 60000
121 | 2020-05-01 | Admin | Manager | Alicia | null | 75000
(3 rows)
cqsh:employee> UPDATE Employee_Info SET Projects = ('Payroll System', 'Attendance Tracker') WHERE Emp_Id = 121;
cqsh:employee> SELECT*FROM Employee_Info;
emp_id | date_of_joining | dept_name | designation | emp_name | projects | salary
-----+-----+-----+-----+-----+-----+
123 | 2019-09-10 | Finance | Analyst | Charlie | null | 50000
122 | 2021-01-15 | IT | Developer | Bob | null | 60000
121 | 2020-05-01 | Admin | Manager | Alicia | ('Attendance Tracker', 'Payroll System') | 75000
(3 rows)
cqsh:employee> ■
Activities Terminal Apr 8 14:24
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~
[qlsh:0] Connected to Test Cluster at 127.0.0.1:9042
[qlsh:0] -> Cassandra 4.1.8 | CQL spec 3.4.6 | Native protocol v5
Use HELP for help.
[qlsh:0]> CREATE KEYSPACE Library
... WITH replication = {'class': 'SimpleStrategy', 'replication_factor': 1};
[qlsh:0]> USE Library;
[qlsh:library]> CREATE TABLE Library_Info (
... Stud_Id int PRIMARY KEY,
... Counter_value Counter,
... Stud_Name text,
... Book_Name text,
... Book_Id int,
... Date_of_Issue date
... );
InvalidRequest: Error from server: code=2200 [Invalid query] message="Cannot mix counter and non counter columns in the same table"
[qlsh:library]> CREATE TABLE Library_Info (
... Stud_Id int PRIMARY KEY,
... Stud_Name text,
... Book_Name text,
... Book_Id int,
... Date_of_Issue date
... );
[qlsh:library]> CREATE TABLE Library_Counter (
... Stud_Id int PRIMARY KEY,
... Counter_value counter
... );
[qlsh:library]> SELECT*FROM Library_Info;
stud_id | book_id | book_name | date_of_issue | stud_name
-----+-----+-----+-----+-----+
(0 rows)
[qlsh:library]> CREATE TABLE Library_Info1(
... Stud_Id int,
... Book_Name text,
... Counter_Value Counter,
... PRIMARY KEY (Stud_Id, Book_Name)
... );
[qlsh:library]> DROP TABLE LIBRARY_INFO;

[qlsh:library]>
[qlsh:library]>
[qlsh:library]> DROP TABLE LIBRARY_INFO1;
[qlsh:library]> CREATE TABLE Library_Info (
... Stud_Id int,
... Book_Name text,
... Counter_Value Counter,
... PRIMARY KEY (Stud_Id, Book_Name)
... );
[qlsh:library]>

```

```

Activities Terminal Apr 8 14:25
bmscsece@bmscsece-HP-Elite-Tower-800-C9-Desktop-PC: ~
[1] 1+ cqlsh:library> BEGIN BATCH
...     UPDATE Library_Info SET Counter_Value = Counter_Value + 1 WHERE Stud_Id = 112 AND Book_Name = 'BDA';
...     UPDATE Library_Info SET Counter_Value = Counter_Value + 1 WHERE Stud_Id = 113 AND Book_Name = 'DSA';
...     UPDATE Library_Info SET Counter_Value = Counter_Value + 1 WHERE Stud_Id = 112 AND Book_Name = 'BDA';
...     APPLY BATCH;
InvalidRequest: Error from server: code=2200 [Invalid query] message="Cannot include a counter statement in a logged batch"
cqlsh:library> BEGIN UNLOGGED BATCH
...     UPDATE Library_Info SET Counter_Value = Counter_Value + 1 WHERE Stud_Id = 112 AND Book_Name = 'BDA';
...     UPDATE Library_Info SET Counter_Value = Counter_Value + 1 WHERE Stud_Id = 113 AND Book_Name = 'DSA';
...     UPDATE Library_Info SET Counter_Value = Counter_Value + 1 WHERE Stud_Id = 112 AND Book_Name = 'BDA';
...     APPLY BATCH;
cqlsh:library> SELECT*FROM Library_Info;
stud_id | book_name | counter_value
-----+
113 | DSA | 1
112 | BDA | 2
(2 rows)
cqlsh:library> SELECT * FROM Library_Info WHERE Stud_Id = 112 AND Book_Name = 'BDA';
stud_id | book_name | counter_value
-----+
112 | BDA | 2
(1 rows)
cqlsh:library> COPY Library_Info TO '/home/your_username/library_info.csv' WITH HEADER = TRUE;
Using 16 child processes
Can't open '/home/your_username/library_info.csv' for writing: [Errno 2] No such file or directory: '/home/your_username/library_info.csv'
cqlsh:library> COPY library_info (stud_id, book_name, counter_value) TO '/home/bmscsece/library_info.csv' WITH HEADER = true;
Using 16 child processes
Starting copy of library.library_info with columns [stud_id, book_name, counter_value].
Processed: 2 rows; Rate: 22 rows/s; Avg. rate: 22 rows/s
2 rows exported to 1 files in 0.105 seconds.
cqlsh:library> COPY library_info (stud_id, book_name, counter_value) FROM '/home/bmscsece/library_info.csv' WITH HEADER = true;
Using 16 child processes
Starting copy of library.library_info with columns [stud_id, book_name, counter_value].
Processed: 1 rows; Rate: 2 rows/s; Avg. rate: 3 rows/s
1 rows imported from 1 files in 0.358 seconds (0 skipped).
cqlsh:library> select*from library.library_info;
stud_id | book_name | counter_value
-----+
113 | DSA | 1
112 | BDA | 4
(2 rows)
cqlsh:library> █

```

EXPERIMENT:4

Perform the following DB operations using Cassandra.

- a) Create a keyspace by name Library
 - b) 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
 - c) Insert the values into the table in batch
 - d) Display the details of the table created and increase the value of the counter
 - e) Write a query to show that a student with id 112 has taken a book “BDA” 2 times.
 - f) Export the created column to a csv file
 - g) Import a given csv dataset from local file system into Cassandra column family

```
vn.max.map.count = 1048575
net.ipv4.tcp_keepalive_time = 300
update-rc.d: warning: start and stop actions are no longer supported; falling back to defaults
bmscse@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~$ cassandra -f
OpenSSL: 64-Bit Server VM Warning: Option UseConMarkSweepC was deprecated in version 9.0 and will likely be removed in a future release.
Command log file: /var/log/cassandra/cassandra.log
Log file: /var/log/cassandra/cassandra.log
[0.000s][error][logging] Error opening log file '/var/log/cassandra/cassandra.log': permission denied.
[0.000s][error][logging] Initialization of output 'file:/var/log/cassandra/cql.log' using options 'filecount=10,filesize=10485760' failed.
Invalid -Xlog option '-Xlog:gc=info,heap+trace,age*=trace,perf*=trace,file=/var/log/cassandra/gc.log:time,uptime,pid,tid,level=filecount=10,filesize=10485760', see error log for details.
ls.
Error: Could not create the Java Virtual Machine.
Error: A fatal exception has occurred. Program will exit.
bmscse@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~$ sudo mkdir -p /var/log/cassandra
sudo: mkdir: directory '/var/log/cassandra' exists
bmscse@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~$ sudo chmod -R 777 /var/log/cassandra
bmscse@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~$ sudo systemctl restart cassandra

bmscse@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~$ ^[[200-sudo nano /etc/cassandra/cassandra-env.sh
sudo: command not found
bmscse@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~$ ~sudo nano /etc/cassandra/cassandra-env.sh
Command 'sudo' not found, did you mean:
  command 'sudo' from deb sudo (1.9.9-9+ubuntu2.4)
  command 'sudo' from deb sudo-ldap (1.9.9-9+ubuntu2.4)
Try: sudo apt install <deb name>
bmscse@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~$ sudo nano /etc/cassandra/cassandra-env.sh
bmscse@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~$ 
bmscse@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~$ sudo systemctl restart cassandra

bmscse@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~$ 
bmscse@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~$ nodetool status
Datacenter: datacenter1
Status=Up/Down
-- Address   Load   Tokens Owns  Host ID  Rack
UN 127.0.0.1 300.28 Kib 16    ?        rack1
Note: Non-system keyspaces don't have the same replication settings, effective ownership information is meaningless
bmscse@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~$ cqlsh
Connected to Test Cluster at 127.0.0.1:9042
[cqlsh 6.1.0 | Cassandra 4.1.8 | CQL spec 3.4.6 | Native protocol v5]
Use HELP for help.
cqlsh> CREATE KEYSPACE Students WITH REPLICATION =
... {'class': 'SimpleStrategy', 'replication_factor': 1};

cqlsh>
cqlsh>
cqlsh>

les Terminal Apr 1 14:25 bmscse@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~
cqsh> USE Students;
cqsh:students> BEGIN BATCH
... INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
... VALUES (1, 'Asha', '2012-03-12', 79.9);
...
... INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
... VALUES (2, 'Krian', '2012-03-12', 89.9);
...
... INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
... VALUES (3, 'Tarun', '2012-03-12', 78.9);
...
... INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
... VALUES (4, 'Samrth', '2012-03-12', 98.9);
...
... INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
... VALUES (5, 'Stintha', '2012-03-12', 67.9);
...
... INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
... VALUES (6, 'Vishwan', '2012-03-12', 56.9);
...
... APPLY BATCH;
InvalidRequest: Error from server: code=2200 [Invalid query] message="table students_info does not exist"
cqsh:students>
cqsh:students> DESCRIBE KEYSPACES;
company      students      system_distributed      system_views
emp          system       system_schema      system_virtual_schema
emp_data     system_auth   system_traces

cqsh:students> DESCRIBE TABLES;
cqsh:students> DESCRIBE TABLES;
cqsh:students> DESCRIBE TABLES;
cqsh:students> CREATE TABLE Students_Info (
...   Roll_No int PRIMARY KEY,
...   StudName text,
...   DateOfJoining timestamp,
...   last_exam_Percent double
... );
cqsh:students> BEGIN BATCH
...   INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
...   VALUES (1, 'Asha', '2012-03-12', 79.9);
...
...   INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
...   VALUES (2, 'Krian', '2012-03-12', 89.9);
...
...   INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
...   VALUES (3, 'Tarun', '2012-03-12', 78.9);
...
...   INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
...   VALUES (4, 'Samrth', '2012-03-12', 98.9);
...
...   INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
...   VALUES (5, 'Stintha', '2012-03-12', 67.9);
...
...   INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
...   VALUES (6, 'Vishwan', '2012-03-12', 56.9);
...
...   APPLY BATCH;
[1]+ Stopped                  cqsh
bmscse@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~$ BEGIN BATCH
INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
VALUES (1, 'Asha', '2012-03-12', 79.9);

INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
VALUES (2, 'Krian', '2012-03-12', 89.9);

INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
VALUES (3, 'Tarun', '2012-03-12', 78.9);

INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
VALUES (4, 'Samrth', '2012-03-12', 98.9);

INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
VALUES (5, 'Stintha', '2012-03-12', 67.9);

INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
VALUES (6, 'Vishwan', '2012-03-12', 56.9);
```

```

les Terminal Apr 1 14:26 bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~
cqlsh:students> BEGIN BATCH
...   INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
...     VALUES (1, 'Asha', '2012-03-12', 79.9);
...
...
...   INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
...     VALUES (2, 'Kiran', '2012-03-12', 89.9);
...
...
...   INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
...     VALUES (3, 'Tarun', '2012-03-12', 78.9);
...
...
...   INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
...     VALUES (4, 'Santh', '2012-03-12', 98.9);
...
...
...   INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
...     VALUES (5, 'Simha', '2012-03-12', 67.9);
...
...
...   INSERT INTO Students_Info (Roll_No, StudName, DateOfJoining, last_exam_Percent)
...     VALUES (6, 'Rohan', '2012-03-12', 56.9);
...
...   APPLY BATCH;
cqlsh:students> SELECT*FROM students_info;
roll_no | dateofjoining | last_exam_percent | studname
-----+-----+-----+-----
  5 | 2012-03-11 18:30:00.000000+0000 |      67.9 | Simha
  1 | 2012-03-11 18:30:00.000000+0000 |      79.9 | Asha
  2 | 2012-03-11 18:30:00.000000+0000 |      89.9 | Kiran
  4 | 2012-03-11 18:30:00.000000+0000 |      98.9 | Santh
  6 | 2012-03-11 18:30:00.000000+0000 |      56.9 | Rohan
  3 | 2012-03-11 18:30:00.000000+0000 |      78.9 | Tarun

(6 rows)
cqlsh:students> SELECT*FROM student_info WHERE roll_no IN(1,2,3);
InvalidRequest: Error from server: code=200 [Invalid query] message="table student_info does not exist"
cqlsh:students> SELECT*FROM student_info WHERE roll_no IN(1,2,3);
InvalidRequest: Error from server: code=200 [Invalid query] message="table student_info does not exist"
cqlsh:students> SELECT*FROM students_info WHERE roll_no IN(1,2,3);
roll_no | dateofjoining | last_exam_percent | studname
-----+-----+-----+-----
  1 | 2012-03-11 18:30:00.000000+0000 |      79.9 | Asha
  2 | 2012-03-11 18:30:00.000000+0000 |      89.9 | Kiran
  3 | 2012-03-11 18:30:00.000000+0000 |      78.9 | Tarun

(3 rows)
cqlsh:students> SELECT*FROM students_info WHERE studname='Asha';
SyntaxException: line 1:47 no viable alternative at input '!' (SELECT*FROM students_info WHERE studname=[ "Asha"]);
cqlsh:students> SELECT * FROM students_info WHERE studname = 'Asha';
InvalidRequest: Error from server: code=200 [Invalid query] message="Cannot execute this query as it might involve data filtering and thus may have unpredictable performance. If you want to execute this query despite the performance unpredictability, use ALLOW FILTERING"
cqlsh:students> SELECT * FROM students_info WHERE studname = 'Asha' ALLOW FILTERING;
roll_no | dateofjoining | last_exam_percent | studname
-----+-----+-----+-----
  1 | 2012-03-11 18:30:00.000000+0000 |      79.9 | Asha

(1 rows)

les Terminal Apr 1 14:27 bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~
cqlsh:students> select roll_no as "USN" from students_info;
USN
-----
  5
  1
  2
  4
  6
  3

(6 rows)
cqlsh:students> UPDATE roll_no as "USN" from students_info;
SyntaxException: line 1:15 mismatched input 'as' expecting K_SET (UPDATE roll_no [as] "US...")
cqlsh:students> ALTER TABLE students_info RENAME roll_no TO USN;
cqlsh:students> select*from students_info;
usn | dateofjoining | last_exam_percent | studname
-----+-----+-----+-----
  5 | 2012-03-11 18:30:00.000000+0000 |      67.9 | Simha
  1 | 2012-03-11 18:30:00.000000+0000 |      79.9 | Asha
  2 | 2012-03-11 18:30:00.000000+0000 |      89.9 | Kiran
  4 | 2012-03-11 18:30:00.000000+0000 |      90.9 | Santh
  6 | 2012-03-11 18:30:00.000000+0000 |      56.9 | Rohan
  3 | 2012-03-11 18:30:00.000000+0000 |      78.9 | Tarun

(6 rows)
cqlsh:students> UPDATE students_info SET studname='David Sheen' WHERE usn=2;
cqlsh:students> select*from students_info;
usn | dateofjoining | last_exam_percent | studname
-----+-----+-----+-----
  5 | 2012-03-11 18:30:00.000000+0000 |      67.9 | Simha
  1 | 2012-03-11 18:30:00.000000+0000 |      79.9 | Asha
  2 | 2012-03-11 18:30:00.000000+0000 |      89.9 | David Sheen
  4 | 2012-03-11 18:30:00.000000+0000 |      90.9 | Santh
  6 | 2012-03-11 18:30:00.000000+0000 |      56.9 | Rohan
  3 | 2012-03-11 18:30:00.000000+0000 |      78.9 | Tarun

(6 rows)
cqlsh:students> UPDATE students_info SET rollno=6 WHERE rollno=3;
InvalidRequest: Error from server: code=200 [Invalid query] message="Undefined column name rollno in table students.students_info"
cqlsh:students> cqlsh:students> UPDATE students_info SET usn=6 WHERE usn=3;
[2]+ Stopped          cqlsh
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC:~$ cqlsh
Connection to localhost Cluster [127.0.0.1:9042]
CQLsh 6.1.0 | Cassandra 4.1.0 | CQL Spec 3.4.6 | Native protocol v5
Use HELP for help.
cqlsh> use students
...
...
...
...
[3]+ Stopped          cqlsh

```

```

les Terminal Apr 1 14:28
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC:~$ cqlsh
[3]+ Stopped                  cqlsh
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC:~$ cqlsh
Connected to Test Cluster at 127.0.0.1:9042
[cqlsh 6.1.0 | Cassandra 4.1.8 | CQL spec 3.4.6 | Native protocol v5]
Use HELP for help.
cqlsh> USE students;
cqlsh> SELECT * FROM students_info;
+-----+-----+-----+
| usn | dateofjoining | last_exam_percent | studname |
+-----+-----+-----+
| 5 | 2012-03-11 18:30:00.000000+0000 | 67.9 | Simha |
| 1 | 2012-03-11 18:30:00.000000+0000 | 79.9 | Asha |
| 2 | 2012-03-11 18:30:00.000000+0000 | 89.9 | David Sheen |
| 4 | 2012-03-11 18:30:00.000000+0000 | 90.9 | Sanarth |
| 6 | 2012-03-11 18:30:00.000000+0000 | 56.9 | Rohan |
| 3 | 2012-03-11 18:30:00.000000+0000 | 78.9 | Tarun |
+-----+
(6 rows)
cqlsh> UPDATE students_info SET usn=6 WHERE usn=3;
InvalidRequest: Error from server: code=200 [Invalid query] message="PRIMARY KEY part usn found in SET part"
cqlsh> UPDATE students_info SET usn=6 WHERE usn=3;
InvalidRequest: Error from server: code=200 [Invalid query] message="PRIMARY KEY part usn found in SET part"
cqlsh> UPDATE students_info SET usn = 6 WHERE usn= 3;
InvalidRequest: Error from server: code=200 [Invalid query] message="PRIMARY KEY part usn found in SET part"
cqlsh> DELETE FROM students_info WHERE usn = 3;
cqlsh> INSERT INTO students_info (usn, dateofjoining, last_exam_percent, studname)
VALUES (6, '2012-03-11 18:30:00', 78.9, 'Tarun');
cqlsh> select*from STUDENTS_INFO;
+-----+-----+-----+
| usn | dateofjoining | last_exam_percent | studname |
+-----+-----+-----+
| 5 | 2012-03-11 18:30:00.000000+0000 | 67.9 | Simha |
| 1 | 2012-03-11 18:30:00.000000+0000 | 79.9 | Asha |
| 2 | 2012-03-11 18:30:00.000000+0000 | 89.9 | David Sheen |
| 4 | 2012-03-11 18:30:00.000000+0000 | 90.9 | Sanarth |
| 6 | 2012-03-11 18:30:00.000000+0000 | 78.9 | Tarun |
+-----+
(6 rows)
cqlsh> DELETE last_exam_percent from students_info where usn=2;
cqlsh> select*from students_info;
+-----+-----+-----+
| usn | dateofjoining | last_exam_percent | studname |
+-----+-----+-----+
| 5 | 2012-03-11 18:30:00.000000+0000 | 67.9 | Simha |
| 1 | 2012-03-11 18:30:00.000000+0000 | 79.9 | Asha |
| 2 | 2012-03-11 18:30:00.000000+0000 | null | David Sheen |
| 4 | 2012-03-11 18:30:00.000000+0000 | 90.9 | Sanarth |
| 6 | 2012-03-11 18:30:00.000000+0000 | 78.9 | Tarun |
+-----+
(5 rows)
cqlsh> DELETE FROM students_info where usn=2;
cqlsh> SELECT*FROM STUDENTS_INFO;
+-----+-----+-----+
| usn | dateofjoining | last_exam_percent | studname |
+-----+-----+-----+
| 5 | 2012-03-11 18:30:00.000000+0000 | 67.9 | Simha |
+-----+
(5 rows)
les Terminal Apr 1 14:29
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC:~$ cqlsh
[5]+ Stopped                  cqlsh
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC:~$ cqlsh
Connected to Test Cluster at 127.0.0.1:9042
[cqlsh 6.1.0 | Cassandra 4.1.8 | CQL spec 3.4.6 | Native protocol v5]
Use HELP for help.
cqlsh> ALTER TABLE STUDENTS_INFO ADD hobbies set<text>;
...
...
Statements are terminated with a ';'. You can press CTRL-C to cancel an incomplete statement.
cqlsh> ALTER TABLE STUDENTS_INFO ADD hobbies set<text>;
...
cqlsh> ALTER TABLE STUDENTS_INFO ADD hobbies set<text>;
cqlsh> select*from students_info;
+-----+-----+-----+-----+
| usn | dateofjoining | hobbies | last_exam_percent | studname |
+-----+-----+-----+-----+
| 5 | 2012-03-11 18:30:00.000000+0000 | null | 67.9 | Simha |
| 1 | 2012-03-11 18:30:00.000000+0000 | null | 79.9 | Asha |
| 2 | 2012-03-11 18:30:00.000000+0000 | null | 89.9 | David Sheen |
| 4 | 2012-03-11 18:30:00.000000+0000 | null | 90.9 | Sanarth |
| 6 | 2012-03-11 18:30:00.000000+0000 | null | 78.9 | Tarun |
+-----+
(5 rows)
cqlsh> ALTER TABLE students_info ADD language list<text>;
cqlsh> select*from students_info;
+-----+-----+-----+-----+
| usn | dateofjoining | hobbies | language | last_exam_percent | studname |
+-----+-----+-----+-----+
| 5 | 2012-03-11 18:30:00.000000+0000 | null | null | 67.9 | Simha |
| 1 | 2012-03-11 18:30:00.000000+0000 | null | null | 79.9 | Asha |
| 2 | 2012-03-11 18:30:00.000000+0000 | null | null | 89.9 | David Sheen |
| 4 | 2012-03-11 18:30:00.000000+0000 | null | null | 90.9 | Sanarth |
| 6 | 2012-03-11 18:30:00.000000+0000 | null | null | 78.9 | Tarun |
+-----+
(5 rows)

```

```

les Terminal Apr 1 14:29
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~

cqlsh:students> select*from students_info;
+-----+-----+-----+-----+
| usn | dateofjoining | hobbies | last_exam_percent | studname |
+-----+-----+-----+-----+
| 5 | 2012-03-11 18:30:00.000000+0000 | null | 67.9 | Srinha |
| 1 | 2012-03-11 18:30:00.000000+0000 | null | 79.9 | Asha |
| 4 | 2012-03-11 18:30:00.000000+0000 | null | 90.9 | Sanarth |
| 6 | 2012-03-11 13:00:00.000000+0000 | null | 78.9 | Tarun |
+-----+-----+-----+-----+
(4 rows)

cqlsh:students> ALTER TABLE students_info ADD language list<text>;
cqlsh:students> select*from students_info;
+-----+-----+-----+-----+
| usn | dateofjoining | hobbies | language | last_exam_percent | studname |
+-----+-----+-----+-----+
| 5 | 2012-03-11 18:30:00.000000+0000 | null | null | 67.9 | Srinha |
| 1 | 2012-03-11 18:30:00.000000+0000 | null | null | 79.9 | Asha |
| 4 | 2012-03-11 18:30:00.000000+0000 | null | null | 90.9 | Sanarth |
| 6 | 2012-03-11 13:00:00.000000+0000 | null | null | 78.9 | Tarun |
+-----+-----+-----+-----+
(4 rows)

cqlsh:students> UPDATE students_info SET hobbies=hobbies+['Chess,Table Tennis'] where usn=1;
cqlsh:students> SELECT*FROM STUDENTS_INFO;
+-----+-----+-----+-----+
| usn | dateofjoining | hobbies | language | last_exam_percent | studname |
+-----+-----+-----+-----+
| 5 | 2012-03-11 18:30:00.000000+0000 | null | null | 67.9 | Srinha |
| 1 | 2012-03-11 18:30:00.000000+0000 | ['Chess,Table Tennis'] | null | 79.9 | Asha |
| 4 | 2012-03-11 18:30:00.000000+0000 | null | null | 90.9 | Sanarth |
| 6 | 2012-03-11 13:00:00.000000+0000 | null | null | 78.9 | Tarun |
+-----+-----+-----+-----+
(4 rows)

cqlsh:students> SELECT*FROM student_info where usn=1;
InvalidRequest: Error from server: code=200 [Invalid query] message="table student_info does not exist"
cqlsh:students> select*from students_info where usn=1;
+-----+-----+-----+-----+
| usn | dateofjoining | hobbies | language | last_exam_percent | studname |
+-----+-----+-----+-----+
| 1 | 2012-03-11 18:30:00.000000+0000 | ['Chess,Table Tennis'] | null | 79.9 | Asha |
+-----+-----+-----+-----+
(1 rows)

cqlsh:students> update students_info set language=language+['Hindi,English'] where usn=1;
cqlsh:students> select*from students_info;
+-----+-----+-----+-----+
| usn | dateofjoining | hobbies | language | last_exam_percent | studname |
+-----+-----+-----+-----+
| 5 | 2012-03-11 18:30:00.000000+0000 | null | null | 67.9 | Srinha |
| 1 | 2012-03-11 18:30:00.000000+0000 | ['Chess,Table Tennis'] | ['Hindi,English'] | 79.9 | Asha |
| 4 | 2012-03-11 18:30:00.000000+0000 | null | null | 90.9 | Sanarth |
| 6 | 2012-03-11 13:00:00.000000+0000 | null | null | 78.9 | Tarun |
+-----+-----+-----+-----+
(4 rows)

cqlsh:students> 
```

```

Activities Terminal Apr 1 14:40
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~

+-----+-----+-----+
| book_name | stud_name | counter_value |
+-----+-----+-----+
(0 rows)

cqlsh:students> BEGIN BATCH
... Inserting 4 records into library_book table
... INSERT INTO library.book (counter_value, book_name, stud_name)
... VALUES (0, 'Big Data Analytics', 'Jeet');
...
... INSERT INTO library.book (counter_value, book_name, stud_name)
... VALUES (0, 'Machine Learning', 'Amit');
...
... INSERT INTO library.book (counter_value, book_name, stud_name)
... VALUES (0, 'Data Science Essentials', 'Ravi');
...
... INSERT INTO library.book (counter_value, book_name, stud_name)
... VALUES (0, 'Cloud Computing', 'Priya');
... APPLY BATCH;
InvalidRequest: Error from server: code=2200 [Invalid query] message="INSERT statements are not allowed on counter tables, use UPDATE instead"
cqlsh:students> BEGIN BATCH
... Inserting 4 records into library_book table with counter initialized to 0
... INSERT INTO library.book (counter_value, book_name, stud_name)
... VALUES (0, 'Big Data Analytics', 'Jeet');
...
... INSERT INTO library.book (counter_value, book_name, stud_name)
... VALUES (0, 'Machine Learning', 'Amit');
...
... INSERT INTO library.book (counter_value, book_name, stud_name)
... VALUES (0, 'Data Science Essentials', 'Ravi');
...
... INSERT INTO library.book (counter_value, book_name, stud_name)
... VALUES (0, 'Cloud Computing', 'Priya');
... APPLY BATCH;
InvalidRequest: Error from server: code=2200 [Invalid query] message="INSERT statements are not allowed on counter tables, use UPDATE instead"
cqlsh:students> UPDATE library.book
... SET counter_value = counter_value + 1
... WHERE book_name = 'Big Data Analytics'
... AND stud_name = 'jeet';
cqlsh:students> SELECT*FROM library_book;

+-----+-----+-----+
| book_name | stud_name | counter_value |
+-----+-----+-----+
| Big Data Analytics | jeet | 1 |
+-----+-----+-----+
(1 rows)

cqlsh:students> 
```

```
Activities Terminal Apr 1 14:41 bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~
...
... INSERT INTO library_book (counter_value, book_name, stud_name)
... VALUES (0, 'Data Science Essentials', 'Ravi');
...
... INSERT INTO library_book (counter_value, book_name, stud_name)
... VALUES (0, 'Cloud Computing', 'Pritya');
... APPLY BATCH;
InvalidRequest: Error from server: code=2200 [Invalid query] message="INSERT statements are not allowed on counter tables, use UPDATE instead"
cqish:students> -- Load data into the counter column
cqish:students> UPDATE library_book
... SET counter_value = counter_value + 1
... WHERE book_name = 'Big data Analytics'
... AND stud_name = 'jeet';
cqish:students> SELECT*FROM library_book;
book_name | stud_name | counter_value
Big data Analytics | jeet | 1
(1 rows)
cqish:students> -- TIME TO LIVE (TTL) Example
cqish:students> CREATE TABLE userlogin (
...     userid int PRIMARY KEY,
...     password text
... );
cqish:students>
cqish:students>
cqish:students> SELECT*FROM userlogin;
userid | password
-----+-----
(0 rows)
cqish:students> INSERT INTO userlogin(userid, password)
... VALUES (1, 'Infy')
... USING TTL 30;
cqish:students> SELECT*FROM userlogin;
userid | password
-----+-----
1 | Infy
(1 rows)
cqish:students> SELECT TTL(password)
... FROM userlogin
... WHERE userid = 1;
ttl(password)
-----
(0 rows)
cqish:students>
```

```
Activities Terminal Apr 1 14:49
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~

CREATE TABLE students.elearninglists (
    id uuid PRIMARY KEY,
    course_id uuid,
    course_order int,
    courseowner text,
    title text
) WITH additional_write_policy = '99p'
    AND bloom_filter_fp_chance = 0.01
    AND caching = {'keys': 'ALL', 'rows_per_partition': 'NONE'}
    AND cdc = false
    AND comment = ''
    AND compression = ['class': 'org.apache.cassandra.db.compaction.SizeTieredCompactionStrategy', 'max_threshold': '32', 'min_threshold': '4']
    AND compression_type = 'GZIP'
    AND crc_check_chance = 1.0
    AND default_time_to_live = 0
    AND extended_schema = false
    AND gc_grace_seconds = 864000
    AND max_index_interval = 2048
    AND memtable_flush_period_in_ms = 0
    AND min_index_interval = 128
    AND read_repair = 'BLOCKING'
    AND speculative_retry = '99p';
cqlsh:students> COPY elearninglists(id, course_order, course_id, courseowner, title)
... TO 'd:/elearninglists.csv' WITH HEADER = TRUE;
Using 16 child processes
Can't open 'd:/elearninglists.csv' for writing: [Errno 2] No such file or directory: 'd:/elearninglists.csv'
cqlsh:students> COPY elearninglists(id, course_order, course_id, courseowner, title)
... TO '/home/bmscsece/Desktop/elearninglists.csv' WITH HEADER = TRUE;
Using 16 child processes

Starting copy of students.elearninglists with columns [id, course_order, course_id, courseowner, title].
Processed: 0 rows; Rate: 0 rows/s; Avg. rate: 0 rows/s
0 rows exported to files in 0.133 seconds.
cqlsh:students> BEGIN BATCH
...     INSERT INTO elearninglists (id, course_order, course_id, courseowner, title)
...         VALUES (uuid(), 1, uuid(), 'John Doe', 'Introduction to Python');
...
...     INSERT INTO elearninglists (id, course_order, course_id, courseowner, title)
...         VALUES (uuid(), 2, uuid(), 'Jane Smith', 'Machine Learning Basics');
...
...     INSERT INTO elearninglists (id, course_order, course_id, courseowner, title)
...         VALUES (uuid(), 3, uuid(), 'Alice Johnson', 'Data Structures and Algorithms ');
...     APPLY BATCH;
cqlsh:students> select*from elearninglists;
          id          | course_id | course_order | courseowner | title
-----+-----+-----+-----+-----+
f6cb9530-c3c6-4619-b26a-7f1f14475290 | 10cedd27-2d81-4481-ae73-d156a2bd0ef0 | 3 | Alice Johnson | Data Structures and Algorithms
f9892a5c-a6e2-4156-a496-2f62e0f77962 | 5b5bbff5-55bd-4726-9460-786a98838883 | 2 | Jane Smith | Machine Learning Basics
59ca570d-0253-440b-b137-003f51ff1f1 | 5d5c2b42-97ef-483d-97ba-33e201ef7b36 | 1 | John Doe | Introduction to Python
cqlsh:students> select*from elearninglists;
          id          | course_id | course_order | courseowner | title
-----+-----+-----+-----+-----+
f6cb9530-c3c6-4619-b26a-7f1f14475290 | 10cedd27-2d81-4481-ae73-d156a2bd0ef0 | 3 | Alice Johnson | Data Structures and Algorithms
f9892a5c-a6e2-4156-a496-2f62e0f77962 | 5b5bbff5-55bd-4726-9460-786a98838883 | 2 | Jane Smith | Machine Learning Basics
59ca570d-0253-440b-b137-003f51ff1f1 | 5d5c2b42-97ef-483d-97ba-33e201ef7b36 | 1 | John Doe | Introduction to Python
cqlsh:students> COPY elearninglists(id, course_order, course_id, courseowner, title)
... TO '/home/bmscsece/Desktop/elearninglists.csv' WITH HEADER = TRUE;
```

```
Activities Terminal Apr 1 14:50
bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~

cqsh:students> COPY elearninglists(id, course_order, course_id, courseowner, title)
... TO 'd:/elearninglists.csv' WITH HEADER = TRUE;
Using 16 child processes
A [3 rows] can't open 'd:/elearninglists.csv' for writing: [Errno 2] No such file or directory: 'd:/elearninglists.csv'
cqsh:students> COPY elearninglists(id, course_order, course_id, courseowner, title)
... TO '/home/bmscsece/Desktop/elearninglists.csv' WITH HEADER = TRUE;
Using 16 child processes
? [3 rows] Starting copy of students.elearninglists with columns [id, course_order, course_id, courseowner, title].
Processed: 0 rows/s Avg. rate: 0 rows/s
0 rows exported to 1 files in 0.133 seconds.
cqsh:students> BEGIN BATCH
...     INSERT INTO elearninglists (id, course_order, course_id, courseowner, title)
...     VALUES (uuid(), 1, uuid(), 'John Doe', 'Introduction to Python');
...     INSERT INTO elearninglists (id, course_order, course_id, courseowner, title)
...     VALUES (uuid(), 2, uuid(), 'Jane Smith', 'Machine Learning Basics');
...     INSERT INTO elearninglists (id, course_order, course_id, courseowner, title)
...     VALUES (uuid(), 3, uuid(), 'Alice Johnson', 'Data Structures and Algorithms');
...     APPLY BATCH;
cqsh:students> select*from elearninglists;
+-----+-----+-----+-----+
| id   | course_id | course_order | courseowner | title
+-----+-----+-----+-----+
| fcc89530-c3c6-4e19-b26a-7f1f14d75290 | 10cded27-2d81-4481-ae73-d156a2b01a0f | 3 | Alice Johnson | Data Structures and Algorithms
| f9892ac5-a6e2-4156-a996-2f52e077962 | 5bb5bbff5-55bb-4726-9460-786a98838883 | 2 | Jane Smith | Machine Learning Basics
| 59ca570d-0253-440b-b137-003f51ff31f1 | 5d5c2b42-97ef-483d-97ba-33e2fe7b36 | 1 | John Doe | Introduction to Python
+-----+-----+-----+-----+
(3 rows)
cqsh:students> COPY elearninglists(id, course_order, course_id, courseowner, title)
... TO '/home/bmscsece/Desktop/elearninglists.csv' WITH HEADER = TRUE;
Using 16 child processes
SSD [3 rows] Starting copy of students.elearninglists with columns [id, course_order, course_id, courseowner, title].
Processed: 3 rows; Rate: 54 rows/s Avg. rate: 54 rows/s
3 rows exported to 1 files in 0.070 seconds.
cqsh:students> COPY elearninglists(id, course_order, course_id, courseowner, title)
... FROM '/home/bmscsece/Desktop/elearninglists.csv' WITH HEADER = TRUE;
Using 16 child processes
Starting copy of students.elearninglists with columns [id, course_order, course_id, courseowner, title].
Processed: 3 rows; Rate: 4 rows/s Avg. rate: 7 rows/s
3 rows imported from 1 files in 0.452 seconds (0 skipped).
cqsh:students> cat persons.csv | cqsh -e "COPY persons(id, fname, lname) FROM STDIN WITH HEADER = TRUE;"  
Invalid syntax near character position 1.  
    cat persons.csv | cqsh -e "COPY persons(id, fname, lname) FROM STDIN WITH HEADER = TRUE;"  
cqsh:students> COPY elearninglists(id, course_order, course_id, courseowner, title)
... TO STDOUT WITH HEADER = TRUE;
+-----+-----+-----+-----+
| id   | course_order | course_id | courseowner | title
+-----+-----+-----+-----+
| fcc89530-c3c6-4e19-b26a-7f1f14d75290, 3,10cded27-2d81-4481-ae73-d156a2b01a0f,Alice Johnson,Data Structures and Algorithms
| f9892ac5-a6e2-4156-a996-2f52e077962,2,5bb5bbff5-55bb-4726-9460-786a98838883,Jane Smith,Machine Learning Basics
| 59ca570d-0253-440b-b137-003f51ff31f1,1,5d5c2b42-97ef-483d-97ba-33e2fe7b36,John Doe,Introduction to Python
+-----+-----+-----+-----+
(3 rows)
cqsh:students>
```

EXPERIMENT-5

Execution of HDFS Commands for interaction with Hadoop Environment. (Minimum 10 commands to be executed)

```
hadoop@bmsecese-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/bin
hadoop@bmsecese-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -cat /abc/file2.txt
hi how are you
how is your job
how is your family
how is your brother
how is your sister

hadoop@bmsecese-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -get /abc/WC.txt /home/hduser/Downloads/WC.txt
get: '/home/hduser/Downloads/WC.txt': No such file or directory: file:///home/hduser/Downloads/WC.txt'
hadoop@bmsecese-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ ls /home/hduser/Downloads/
ls: cannot access '/home/hduser/Downloads/': No such file or directory
hadoop@bmsecese-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ mkdir -p /home/hduser/Downloads/
mkdir: cannot create directory '/home/hduser/': Permission denied
hadoop@bmsecese-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ sudo mkdir -p /home/hduser/Downloads/
[sudo] password for hadoop:
hadoop@bmsecese-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ ls /home/hduser/
ls: /home/hduser: No such file or directory

hadoop@bmsecese-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -get /abc/WC.txt /home/hduser/Downloads/WC.txt
get: '/abc/WC.txt': No such file or directory
hadoop@bmsecese-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -get /abc/WC.txt /home/hduser/Downloads/WC.txt
get: '/abc/WC.txt': No such file or directory
hadoop@bmsecese-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -ls abc/
Found 2 items
-rw-r--r-- 1 hadoop supergroup 90 2025-04-15 14:34 /abc/file1.txt
-rw-r--r-- 1 hadoop supergroup 90 2025-04-15 14:35 /abc/file2.txt
hadoop@bmsecese-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -get /abc/file1.txt /home/hduser/Downloads/file1.txt
hadoop@bmsecese-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ rm -l /home/hduser/Downloads/file1.txt
drwxr-xr-x 2 root root 4096 Apr 15 14:37 /home/hduser/downloads/
hadoop@bmsecese-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ sudo chown -R hadoop:hadoop /home/hduser/Downloads/
hadoop@bmsecese-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ ls -l /home/hduser/Downloads/
drwxr-xr-x 2 hadoop hadoop 4096 Apr 15 14:37 /home/hduser/downloads/
hadoop@bmsecese-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -get /abc/file2.txt /home/hduser/Downloads/
hadoop@bmsecese-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -getmerge /abc/WC.txt /abc/MC2.txt /home/hduser/Desktop/Merge.txt
getmerge: '/abc/WC.txt': No such file or directory
getmerge: '/abc/MC2.txt': No such file or directory
hadoop@bmsecese-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -getmerge /abc/file1.txt /abc/file2.txt /home/hduser/Desktop/Merge.txt
getmerge: Mkdirs failed to create file:/home/hduser/Desktop (exists already, cwd=file:/home/hduser/hadoop/sbin)
hadoop@bmsecese-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ sudo mkdir -p /home/hduser/Desktop
hadoop@bmsecese-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ sudo chown -R hadoop:hadoop /home/hduser/Desktop
sudo chmod 775 /home/hduser/Desktop
hadoop@bmsecese-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -getmerge /abc/file1.txt /abc/file2.txt /home/hduser/Desktop/Merge.txt
hadoop@bmsecese-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ cat /home/hduser/Desktop/Merge.txt
hi how are you
how is your job
how is your family
how is your brother
how is your sister

hi how are you
how is your job
how is your family
how is your brother
how is your sister
```

```
hadoop@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/bin

hadoop@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -copyToLocal /abc/WC.txt /home/hduser/Desktop
copyToLocal: '/abc/WC.txt': No such file or directory
hadoop@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -copyToLocal /abc/file1.txt /home/hduser/Desktop
hadoop@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ cat /home/hduser/Desktop/file1.txt
hi how are you
how is your job
how is your family
how is your brother
how is your sister

hadoop@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -cat /abc/file1.txt
hi how are you
how is your job
how is your family
how is your brother
how is your sister

hadoop@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -cat /abc/file2.txt
hi how are you
how is your job
how is your family
how is your brother
how is your sister

hadoop@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -cat /abc/file2.txt
hi how are you
how is your job
how is your family
how is your brother
how is your sister

hadoop@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -mv /abc /FFF
hadoop@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -ls /FFF
Found 2 items
-rw-r--r-- 1 hadoop supergroup 98 2025-04-15 14:34 /FFF/file1.txt
-rw-r--r-- 1 hadoop supergroup 98 2025-04-15 14:35 /FFF/file2.txt
hadoop@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ 9. Cp Hadoop HDFS

cp Command Usage cp Hadoop HDFS

cp Command Example

hadoop fs -cp /CSE/ /LLL
hadoop fs -ls /LLL

The cp command copies a file from one directory to another directory within the HDFS.

9. command not found
cp: target 'HDFS' is not a directory
cp: cannot stat 'Command: No such file or directory
cp: '/CSE/': No such file or directory
ls: '/LLL': No such file or directory
Command 'The' not found; did you mean:
  command 'the' from deb tar (3.3-rc1-3ubulld1)
Try: sudo apt install <deb name>
HDFS: not found

hadoop@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hadoop fs -cp /abc/file1.txt /newDir/
cp: '/newDir/' is not a file or directory
hadoop@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hadoop fs -mkdir /destinationDir
hadoop@bmscse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hadoop fs -cp /abc/file1.txt /destinationDir/
```

```

hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC: ~/hadoop/sbin
how is your brother
how is your sister
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -mv /abc /FFF
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -ls /FFF
Found 2 items
-rw-r--r-- 1 hadoop supergroup 90 2025-04-15 14:34 /FFF/file1.txt
-rw-r--r-- 1 hadoop supergroup 90 2025-04-15 14:35 /FFF/file2.txt
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ 9. Cp Hadoop HDFS
cp Command Usage cp Hadoop HDFS
cp Command Example
hadoop fs -cp /CSE/ /LLL
hadoop fs -ls /LLL
A The cp command copies a file from one directory to another directory within the
HDFS.
? 9.: command not found
cp: target '/HDFS' is not a directory
cp: target '/CSE/' is not a file or directory
ls: '/LLL': No such file or directory
Command 'The' not found, did you mean:
  command 'he' from deb node-he (1.2.0-3)
  command 'the' from deb the (3.3-crc3-build1)
Try 'hadoop fs --help' for more information about HDFS names.
HDFS: command not found
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hadoop fs -cp /abc/file1.txt /newDir/
cp: '/newDir/': No such file or directory
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -mkdir /destinationDir
cp: '/abc/file1.txt': No such file or directory
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hadoop fs -cp /abc/file1.txt /destinationDir/
ls: '/abc/': such file or directory
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -ls /abc/
Found 5 items
drwxr-xr-x 1 hadoop supergroup 0 2025-04-15 14:35 /FFF
drwxr-xr-x 1 hadoop supergroup 0 2025-04-15 14:47 /destinationDir
drwxr-xr-x 1 hadoop supergroup 0 2024-05-13 15:12 /newDataFlair
drwxr-xr-x 1 hadoop supergroup 0 2024-05-13 14:36 /oldData
drwxr-xr-x 1 hadoop supergroup 0 2025-04-15 14:19 /rgs
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -ls /FFF
\Found 2 items
-rw-r--r-- 1 hadoop supergroup 90 2025-04-15 14:34 /FFF/file1.txt
-rw-r--r-- 1 hadoop supergroup 90 2025-04-15 14:35 /FFF/file2.txt
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hadoop fs -cp /FFF/file1.txt /destinationDir/
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hadoop fs -cp /FFF/file2.txt /destinationDir/
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -ls /destinationDir/
Found 2 items
-rw-r--r-- 1 hadoop supergroup 90 2025-04-15 14:49 /destinationDir/file1.txt
-rw-r--r-- 1 hadoop supergroup 90 2025-04-15 14:49 /destinationDir/file2.txt
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ 5-5-
```



```

hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC: ~/hadoop/sbin
SUBCOMMAND may print help when invoked w/o parameters or with -h.
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC: $ cd /home/hadoop/hadoop/sbin
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ ./start-dfs.sh
Starting namenodes on [localhost]
localhost: namenode ls running as process 9990. Stop it first and ensure /tmp/hadoop-hadoop-namenode.pid file is empty before retry.
Starting datanodes
localhost: datanode is running as process 10056. Stop it first and ensure /tmp/hadoop-hadoop-datanode.pid file is empty before retry.
Starting secondary namenodes [bmsccse-HP-Elite-Tower-800-G9-Desktop-PC]
bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:secondarynamenode ls running as process 10334. Stop it first and ensure /tmp/hadoop-hadoop-secondarynamenode.pid file is empty before retry.
Starting resourcemanager
resourcemanager is running as process 10644. Stop it first and ensure /tmp/hadoop-hadoop-resourcemanager.pid file is empty before retry.
Starting nodemanagers
localhost: nodemanager is running as process 10807. Stop it first and ensure /tmp/hadoop-hadoop-nodemanager.pid file is empty before retry.
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ jps
13012 Jps
10644 ResourceManager
10807 NodeManager
10056 DataNode
9990 NameNode
10334 SecondaryNameNode
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ ./start-dfs.sh
Starting namenodes on [localhost]
localhost: namenode is running as process 9990. Stop it first and ensure /tmp/hadoop-hadoop-namenode.pid file is empty before retry.
Starting datanodes
localhost: datanode is running as process 10056. Stop it first and ensure /tmp/hadoop-hadoop-datanode.pid file is empty before retry.
Starting secondary namenodes [bmsccse-HP-Elite-Tower-800-G9-Desktop-PC]
bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:secondarynamenode ls running as process 10334. Stop it first and ensure /tmp/hadoop-hadoop-secondarynamenode.pid file is empty before retry.
Starting resourcemanager
resourcemanager is running as process 10644. Stop it first and ensure /tmp/hadoop-hadoop-resourcemanager.pid file is empty before retry.
Starting nodemanagers
localhost: nodemanager is running as process 10807. Stop it first and ensure /tmp/hadoop-hadoop-nodemanager.pid file is empty before retry.
jps
13012 Jps
10644 ResourceManager
10807 NodeManager
10056 DataNode
9990 NameNode
10334 SecondaryNameNode
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hadoop fs -ls /
Found 2 items
drwxr-xr-x 1 hadoop supergroup 0 2024-05-13 15:12 /newDataFlair
drwxr-xr-x 1 hadoop supergroup 0 2024-05-13 14:36 /oldData
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hadoop fs -copyToLocal /home/rekhags/Desktop/file1.txt /rgs/test.txt
copyFromLocal: '/rgs/test.txt': No such file or directory: 'hdfs://localhost:9000/rgs/test.txt'
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hdfs dfs -mkdir /rgs
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hadoop fs -copyFromLocal /home/rekhags/Desktop/file1.txt /rgs/test.txt
copyFromLocal: '/home/rekhags/Desktop/file1.txt': No such file or directory
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ ls /home/hadoop/Desktop/
file1.txt first.txt invited.jad MC_1
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ nano /home/hadoop/Desktop/file1.txt
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hadoop fs -copyFromLocal /home/hadoop/Desktop/file1.txt /rgs/test.txt
hadoop@bmsccse-HP-Elite-Tower-800-G9-Desktop-PC:~/hadoop/sbin$ hadoop fs -mkdir /rgs
```

```
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~[[200-sudo apt update
sudo: command not found
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~$ -sudo apt update
Command not found, did you mean:
  command 'sudo' from deb sudo (1:9.9~ubuntu2.4)
  command 'sudo' from deb sudo-1dep (1:9.9~ubuntu2.4)
Try: sudo apt install <deb name>
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~$ tar -xvf hadoop-3.x.x.tar.gz
tar (child): hadoop-3.x.x.tar.gz: Cannot open: No such file or directory
tar: Child returned status 2
tar: Error is not recoverable: exiting now
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~$ sudo apt update
[sudo] password for hadoop:
Hit: https://packages.microsoft.com/repos/code stable InRelease
Hit: https://deb.debian.org/debian jessie InRelease
Get:3 http://dl.google.com/linux/chrome/deb stable/main amd64 Packages [1,025 B]
Hit:4 http://in.archive.ubuntu.com/ubuntu jammy InRelease
Get:5 http://in.archive.ubuntu.com/ubuntu jammy-updates InRelease [128 kB]
Hit:5 https://repo.mongodb.org/apt/ubuntu jammy/mongodb-org/6.0 InRelease
Get:7 https://dl.google.com/linux/chrome/deb stable/main amd64 Packages [1,214 B]
Hit:8 http://repo.mongodb.org/apt/ubuntu jammy/mongodb-org/7.0 InRelease
Hit:9 http://archive.ubuntu.com/ubuntu jammy-backports InRelease
Hit:8 https://apt.jfrog.org/artifactory/cassandra-deb 4ix InRelease
Fetched 131 KB in 3s (42.9 kB/s)
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
S: /var/lib/dpkg/info/apt.list.upgradable to see them.
W: https://repo.mongodb.org/apt/ubuntu/dists/jammy/mongodb-org/6.0/InRelease: Key is stored in legacy trusted.gpg keyring (/etc/apt/trusted.gpg), see the DEPRECATION section in apt-key(8) for details.
W: https://debian.cassandra.apache.org/dists/4ix/InRelease: Key is stored in legacy trusted.gpg keyring (/etc/apt/trusted.gpg), see the DEPRECATION section in apt-key(8) for details.
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~$ sudo apt install openjdk-11-jdk
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
openjdk-11-jdk is already the newest version (11.0.26+4~ubuntu1~22.04).
The following packages were automatically installed and are no longer required:
  linux-headers-6.2.0-26-generic linux-hwe-6.2.0-26-generic
  linux-image-6.2.0-26-generic linux-modules-6.2.0-26-generic
  linux-modules-extra-6.2.0-26-generic
Use 'sudo apt autoremove' to remove them.
0 upgraded, 0 newly installed, 0 to remove and 344 not upgraded.
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~$ java -version
openjdk version "11.0.26" 2025-01-21
OpenJDK Runtime Environment (build 11.0.26+4~post-Ubuntu-ubuntu122.04)
OpenJDK 64-Bit Server VM (build 11.0.26+4~post-Ubuntu-ubuntu122.04, mixed mode, sharing)
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~$ mkdir -p ~/hadoop_wordcount
Cd ~/hadoop_wordcount
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~/hadoop_wordcount$ nano WordCount.java
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~/hadoop_wordcount$ javac -classpath $(hadoop classpath) -d /home/hadoop/wordcount_classes WordCount.java
WordCount.java:18: error: cannot find symbol
    public static class TokenizerMapper extends MapReduceBase implements Mapper<Object, Text, Text, IntWritable> {
                                         ^
      symbol:  class MapReduceBase
      location: class WordCount
WordCount.java:22: error: cannot find symbol
    WordCount.java:22: error: cannot find symbol
        ^
      symbol:  class MapReduceBase
      location: class WordCount

```

EXPERIMENT-6

Implement Wordcount program on Hadoop framework

Code with Screenshots:

The screenshot shows two terminal sessions on a Linux desktop. The top session, dated April 29 15:30, runs a Hadoop command to copy a local file 'sample.txt' to HDFS and then run a MapReduce job using the 'WC' driver. It outputs logs from the 'MyDriver' application, which includes configuration details like HADOOP_HOME and Hadoop version 2.0.0. The bottom session, dated April 29 15:31, shows the execution of the same command again, but this time it fails with an error code 1 due to a ClassNotFoundException for 'org.apache.hadoop.mapreduce.v2.app.MRAppMaster'. Both sessions show the terminal interface with tabs for Activities and Terminal, and icons for various applications like a browser and file manager.

```
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~$ hadoop fs -mkdir /input
hadoop fs -copyFromLocal /home/hadoop/Desktop/sample.txt /input/
mkdir: /input: File exists
copyFromLocal: /home/hadoop/Desktop/sample.txt: No such file or directory
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~$ hadoop fs -rm -r /input
hadoop fs -mkdir /input
Deleted /input
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~$ ls /home/hadoop/Desktop/
WC.jar
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~$ nano /home/hadoop/Desktop/sample.txt
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~$ hadoop fs -copyFromLocal /home/hadoop/Desktop/sample.txt /input/
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~$ hadoop fs -mkdir /input
mkdir: /input: File exists
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~$ hadoop fs -copyFromLocal /home/hadoop/Desktop/sample.txt /input/
copyFromLocal: /input/sample.txt: File exists
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~$ hadoop fs -ls /input
Found 1 item
-rw-r--r-- 1 hadoop supergroup 89 2025-04-29 15:21 /input/sample.txt
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~$ hadoop jar /home/hadoop/Desktop/WC.jar MyDriver /input/sample.txt /output
2025-04-29 15:22:52,183 INFO client.DefaultNoHARNFailoverProxyProvider: Connecting to ResourceManager at /0.0.0.0:8032
2025-04-29 15:22:52,243 INFO client.DefaultNoHARNFailoverProxyProvider: Connecting to ResourceManager at /0.0.0.0:8032
2025-04-29 15:22:52,345 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/job_1745916871358_0002
2025-04-29 15:22:52,508 INFO mapred.FileInputFormat: Total input files to process : 1
2025-04-29 15:22:52,572 INFO mapreduce.JobsUBL: Submitting tokens for job: job_1745916871358_0002
2025-04-29 15:22:52,670 INFO mapreduce.JobsUBL: Submitting tokens for job: job_1745916871358_0002
2025-04-29 15:22:52,740 INFO mapreduce.JobsUBL: Executing with tokens: []
2025-04-29 15:22:52,740 INFO mapreduce.JobsUBL: Error: Configuration property 'mapred.job.tracker' not found
2025-04-29 15:22:52,740 INFO mapreduce.JobsUBL: Error: Configuration property 'mapred.job.tracker' not found
2025-04-29 15:22:52,756 INFO impl.YarnClientImpl: Submitted application application_1745916871358_0002
2025-04-29 15:22:52,781 INFO mapreduce.Job: The url to track the job: http://bmscsece-HP-Elite-Tower-800-G9-Desktop-PC:8088/proxy/application_1745916871358_0002/
2025-04-29 15:22:55,795 INFO mapreduce.Job: Job job_1745916871358_0002 running in uber mode : false
2025-04-29 15:22:55,802 INFO mapreduce.Job:  map 0% reduce 0%
2025-04-29 15:22:55,802 INFO mapreduce.Job: Job job_1745916871358_0002 failed with state FAILED due to: Application application_1745916871358_0002 failed 2 times due to AM Container for appattempt_1745916871358_0002 exited with exitcode: 1
Failing this attempt.Diagnostics: [2025-04-29 15:22:55.804]Exception From container-launch.
Container id: container_1745916871358_0002_02_000001
Exit code: 1
[2025-04-29 15:22:55.806]Container exited with a non-zero exit code 1. Error file: prelaunch.err.
Last 4096 bytes of prelaunch.err :
Last 4096 bytes of stderr :
Error: Could not find or load main class org.apache.hadoop.mapreduce.v2.app.MRAppMaster
Caused by: java.lang.ClassNotFoundException: org.apache.hadoop.mapreduce.v2.app.MRAppMaster
Please check whether your <HADOOP_HOME>/etc/hadoop/mapred-site.xml contains the below configuration:
<property>
  <name>yarn.app.mapreduce.an.env</name>
  <value>HADOOP_MAPRED_HOME=${full path of your hadoop distribution directory}</value>
</property>
<property>
  <name>mapreduce.map.env</name>
  <value>HADOOP_MAPRED_HOME=${full path of your hadoop distribution directory}</value>
</property>
<property>
  <name>mapreduce.reduce.env</name>
  <value>HADOOP_MAPRED_HOME=${full path of your hadoop distribution directory}</value>
</property>
[2025-04-29 15:22:55.806]Container exited with a non-zero exit code 1. Error file: prelaunch.err.
Last 4096 bytes of prelaunch.err :
Last 4096 bytes of stderr :
Error: Could not find or load main class org.apache.hadoop.mapreduce.v2.app.MRAppMaster
Caused by: java.lang.ClassNotFoundException: org.apache.hadoop.mapreduce.v2.app.MRAppMaster
Please check whether your <HADOOP_HOME>/etc/hadoop/mapred-site.xml contains the below configuration:
<property>
  <name>yarn.app.mapreduce.an.env</name>
  <value>HADOOP_MAPRED_HOME=${full path of your hadoop distribution directory}</value>
</property>
<property>
  <name>mapreduce.map.env</name>
  <value>HADOOP_MAPRED_HOME=${full path of your hadoop distribution directory}</value>
</property>
<property>
  <name>mapreduce.reduce.env</name>
  <value>HADOOP_MAPRED_HOME=${full path of your hadoop distribution directory}</value>
</property>
For more detailed output, check the application tracking page: http://bmscsece-HP-Elite-Tower-800-G9-Desktop-PC:8088/cluster/app/application_1745916871358_0002 Then click on links to logs of each attempt.
Failing the application.
2025-04-29 15:22:55,811 INFO mapreduce.Job: Counters:
Exception in thread "main" java.io.IOException: Job failed!
  at org.apache.hadoop.mapred.JobClient.runJob(JobClient.java:876)
  at MyDriver.run(MyDriver.java:28)
  at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:82)
  at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:97)
  at MyDriver.main(MyDriver.java:33)
  at java.base/jdk.internal.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
  at java.base/jdk.internal.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
  at java.base/jdk.internal.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
  at java.base/java.lang.reflect.Method.invoke(Method.java:566)
  at org.apache.hadoop.util.RunJar.run(RunJar.java:238)
  at org.apache.hadoop.util.RunJar.main(RunJar.java:241)
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~$ hadoop fs -ls /output
ls: '/output': No such file or directory
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~$ nano /home/hadoop/hadoop/etc/hadoop/mapred-site.xml
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~$ startall.sh
WARNING: Attempting to start all Apache Hadoop daemons as hadoop in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
localhost: namenode is running as process 7941. Stop it first and ensure /tmp/hadoop-hadoop-namenode.pid file is empty before retry.
Starting datanodes. 1 datanode is running as process 8118. Stop it first and ensure /tmp/hadoop-hadoop-datanode.pid file is empty before retry.
Starting secondary namenodes [bmscsece-HP-Elite-Tower-800-G9-Desktop-PC]
Starting secondarynamenode is running as process 8480. Stop it first and ensure /tmp/hadoop-hadoop-secondarynamenode.pid file is empty before retry.
Starting resourcemanager
resourcemanager is running as process 8697. Stop it first and ensure /tmp/hadoop-hadoop-resourcemanager.pid file is empty before retry.
```

Activities Terminal Apr 29 15:31 hadoop@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC:~

```

resourceManager is running as process 8687. Stop it first and ensure /tmp/hadoop-hadoop/resourceManager.pid file is empty before retry.
Starting nodemanagers
localhost: nodemanager is running as process 8852. Stop it first and ensure /tmp/hadoop-hadoop-nodemanager.pid file is empty before retry.
hadoop@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC:~$ jps
15538 Jps
8852 NodeManager
7941 NameNode
8118 DataNode
10892 org.eclipse.equinix.launcher_1.6.1000.v20250227-1734.jar
8087 ResourceManager
hadoop@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC:~$ cd ~/Desktop # or navigate to any other directory where you want to create the file
hadoop@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC:~/Desktop$ nano sample.txt
hadoop@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC:~/Desktop$ cat ~
8462 SecondaryNameNode
8052 JobManager
7941 NameNode
8118 DataNode
15593 Jps
10892 org.eclipse.equinix.launcher_1.6.1000.v20250227-1734.jar
8087 ResourceManager
hadoop@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC:~$ hadoop fs -mkdir /rgs
mkdir: /rgs: File exists
hadoop@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC:~$ hadoop fs -copyFromLocal ~/Desktop/sample.txt /rgs/test.txt
copyFromLocal: /rgs/test.txt: File exists
hadoop@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC:~$ hadoop fs -ls /rgs
Found 1 item
-rw-r--r-- 1 hadoop supergroup 90 2025-04-15 14:19 /rgs/test.txt
hadoop@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC:~$ hadoop fs -cat /rgs/test.txt
how are you
how is your job
how is your family
how is your brother
how is your sister

hadoop@bmseccse-HP-Elite-Tower-800-G9-Desktop-PC:~$ hadoop jar /home/hadoop/Desktop/WC.jar MyDriver /rgs /output
2025-04-29 15:29:14,897 INFO impl.MetricsConfig: Loaded properties from hadoop-metrics2.properties
2025-04-29 15:29:14,935 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).
2025-04-29 15:29:14,935 INFO impl.MetricsSystemImpl: Jobtracker metrics system started
2025-04-29 15:29:14,941 INFO impl.MetricsSystemImpl: Jobtracker metrics system already initialized!
2025-04-29 15:29:14,996 WARN mapred.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
2025-04-29 15:29:15,058 INFO mapred.FileInputFormat: Total input files to process : 1
2025-04-29 15:29:15,064 INFO mapred.Jobsubmitter: number of splits:1
2025-04-29 15:29:15,122 INFO mapred.Jobsubmitter: Submitting tokens for job: job_local537170479_0001
2025-04-29 15:29:15,122 INFO mapred.Jobsubmitter: Executing with tokens: []
2025-04-29 15:29:15,183 INFO mapred.Jobsubmitter: Job: Running job: http://localhost:8080/
2025-04-29 15:29:15,183 INFO mapred.Jobsubmitter: Job: Running job: job_local537170479_0001
2025-04-29 15:29:15,188 INFO mapred.LocalJobRunner: OutputCommitter set in config null
2025-04-29 15:29:15,188 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2025-04-29 15:29:15,188 INFO output.FileOutputCommitter: File Output Committer skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
2025-04-29 15:29:15,221 INFO mapred.LocalJobRunner: Waiting for map tasks
2025-04-29 15:29:15,222 INFO mapred.LocalJobRunner: Starting task: attempt_local537170479_0001_m_000000_0
2025-04-29 15:29:15,222 INFO mapred.LocalJobRunner: OutputCommitter set in config null
2025-04-29 15:29:15,222 INFO mapred.LocalJobRunner: OutputCommitter: org.apache.hadoop.mapred.FileOutputCommitter
2025-04-29 15:29:15,222 INFO mapred.LocalJobRunner: OutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
2025-04-29 15:29:15,222 INFO mapred.LocalJobRunner: Waiting for map tasks
2025-04-29 15:29:15,222 INFO mapred.LocalJobRunner: Starting task: attempt_local537170479_0001_m_000000_0
2025-04-29 15:29:15,231 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2025-04-29 15:29:15,231 INFO output.FileOutputCommitter: File Output Committer skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
2025-04-29 15:29:15,239 INFO mapred.Task: Using ResourceCalculatorProcessTree: []
2025-04-29 15:29:15,242 INFO mapred.MapTask: Processing split: hdfs://localhost:9000/rgs/test.txt:0+90
2025-04-29 15:29:15,252 INFO mapred.MapTask: numReduceTasks: 1
2025-04-29 15:29:15,281 INFO mapred.MapTask: EQUATOR: 0 kvl 26214396(104857584)
2025-04-29 15:29:15,281 INFO mapred.MapTask: mapreduce.task.io.sort.mb: 100
2025-04-29 15:29:15,281 INFO mapred.MapTask: mapreduce.map.memory.buffer.size = 104857600
2025-04-29 15:29:15,288 INFO mapred.MapTask: kvstart = 26214396(104857584); kvend = 26214320(104857280); length = 77/6553600
2025-04-29 15:29:15,293 INFO mapred.MapTask: Map output collector class = org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2025-04-29 15:29:15,311 INFO mapred.LocalJobRunner:
File System Counters
FILE: Number of bytes read=4156
FILE: Number of bytes written=642471
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=90
HDFS: Number of bytes written=0
HDFS: Number of read operations=5
HDFS: Number of large read operations=0
HDFS: Number of write operations=1
HDFS: Number of bytes read erasure-coded=0
Map Reduce framework
Map input records=2
Map output records=2
Map output bytes=169
Map output materialized bytes=215

```

Activities Terminal Apr 29 15:32

```

hadoop@bmsece-HP-Elite-Tower-800-G9-Desktop-PC: ~

2025-04-29 15:29:15,316 INFO mapred.MapTask: burststart = 0; butend = 169; butvold = 104857600
2025-04-29 15:29:15,316 INFO mapred.MapTask: burststart = 26214396(104857584); kvend = 26214320(104857280); length = 77/6553600
2025-04-29 15:29:15,325 INFO mapred.Task: Task attempt_local537170479_0001_m_000000_0 split=1
2025-04-29 15:29:15,325 INFO mapred.Task: Task attempt_local537170479_0001_m_000000_0 ls done. And is in the process of committing
2025-04-29 15:29:15,327 INFO mapred.LocalJobRunner: hdfs://localhost:9000/test.txt:0:99
2025-04-29 15:29:15,327 INFO mapred.Task: Task 'attempt_local537170479_0001_m_000000_0' done.
2025-04-29 15:29:15,329 INFO mapred.Task: Final Counters for attempt_local537170479_0001_m_000000_0: Counters: 23
File System Counters
FILE: Number of bytes read=4169
FILE: Number of bytes written=642471
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=96
HDFS: Number of bytes written=96
HDFS: Number of read operations=5
HDFS: Number of large read operations=0
HDFS: Number of write operations=1
HDFS: Number of bytes read erasure-coded=0
Map-Reduce Framework
Map input records=2
Map output records=20
Map output bytes=169
Map output materialized bytes=215
Input split bytes=86
Combine input records=0
Spilled Records=20
Shuffled Maps=0
Merged Map outputs=0
GC time elapsed (ms)=0
Total committed heap usage (bytes)=633339904
File Input Format Counters
Bytes Read=0
2025-04-29 15:29:15,330 INFO mapred.LocalJobRunner: Finishing task: attempt_local537170479_0001_m_000000_0
2025-04-29 15:29:15,330 INFO mapred.LocalJobRunner: map task executor complete.
2025-04-29 15:29:15,331 INFO mapred.LocalJobRunner: Waiting for reduce tasks
2025-04-29 15:29:15,331 INFO mapred.LocalJobRunner: Starting task: attempt_local537170479_0001_r_000000_0
2025-04-29 15:29:15,332 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2025-04-29 15:29:15,335 INFO mapred.Task: Using ResourceCalculatorProcessTree : []
2025-04-29 15:29:15,335 INFO mapred.ReduceTask: Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@27a9203f
2025-04-29 15:29:15,337 WARN org.apache.hadoop.metrics.sink.MetricSystemImpl: Jobtracker metric system already initialized!
2025-04-29 15:29:15,341 INFO reduce.EventFetcher: attempt_local537170479_0001_r_000000_0 Thread started: EventFetcher for fetching Map Completion Events
2025-04-29 15:29:15,345 INFO mapred.Task: Using ResourceCalculatorProcessTree : []
2025-04-29 15:29:15,347 INFO reduce.InMemoryMapOutput: Read 211 bytes from map-output for attempt_local537170479_0001_r_000000_0 decomp: 211 len: 215 to MEMORY
2025-04-29 15:29:15,358 INFO reduce.MergeManagerImpl: closeInMemoryFile => map-output of size: 211, InMemoryMapOutputs.size() => 1, commitMemory -> 0, usedMemory -> 211
2025-04-29 15:29:15,358 INFO reduce.EventFetcher: EventFetcher is interrupted.. Returning
2025-04-29 15:29:15,359 INFO mapred.LocalJobRunner: 1 / 1 copied.
2025-04-29 15:29:15,359 INFO mapred.MergerManagerImpl: finalMerge called with 1 in-memory map-outputs and 0 on-disk map-outputs
2025-04-29 15:29:15,361 INFO mapred.Merger: Merged 1 segments
2025-04-29 15:29:15,362 INFO mapred.Merger: Down to the last merge-pass, with 1 segments left of total size: 205 bytes
2025-04-29 15:29:15,362 INFO mapred.MergerManagerImpl: Merged 1 segments, 211 bytes to disk to satisfy reduce memory limit
2025-04-29 15:29:15,363 INFO mapred.Merger: Merging 1 files, 215 bytes from disk
2025-04-29 15:29:15,363 INFO reduce.MergeManagerImpl: Merging 1 files, 0 bytes from memory into reduce
2025-04-29 15:29:15,363 INFO reduce.MergeManagerImpl: Merging 0 sorted segments, 0 bytes from memory into reduce
2025-04-29 15:29:15,363 INFO mapred.Merger: Merging 1 sorted segments

```

Activities Terminal Apr 29 15:33

```

hadoop@bmsece-HP-Elite-Tower-800-G9-Desktop-PC: ~

Bytes Read=90
2025-04-29 15:29:15,329 INFO mapred.local.Jobrunner: Finishing task: attempt_local537170479_0001_m_000000_0
2025-04-29 15:29:15,329 INFO mapred.local.Jobrunner: map task executor complete.
2025-04-29 15:29:15,331 INFO mapred.localJobrunner: Waiting for reduce tasks
2025-04-29 15:29:15,331 INFO mapred.localJobrunner: Starting task: attempt_local537170479_0001_r_000000_0
2025-04-29 15:29:15,335 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2025-04-29 15:29:15,335 INFO output.FileOutputCommitter: File Output Committer skip cleanup _temporary folders under output directory=false, ignore cleanup failures: false
2025-04-29 15:29:15,335 INFO mapred.Task: Using ResourceCalculatorProcessTree : []
2025-04-29 15:29:15,337 WARN org.apache.hadoop.metrics.sink.MetricSystemImpl: Jobtracker metric system already initialized!
2025-04-29 15:29:15,344 INFO reduce.MergeManagerImpl: MergerManager: memoryLmt=1456996352, mergeThreshold=3846470400, ioSortFactor=10, memToMemMergeOutputsThreshold=10
2025-04-29 15:29:15,344 INFO reduce.EventFetcher: attempt_local537170479_0001_r_000000_0 Thread started: EventFetcher for fetching Map Completion Events
2025-04-29 15:29:15,357 INFO reduce.InMemoryMapOutput: Read 211 bytes from map-output for attempt_local537170479_0001_r_000000_0 decomp: 211 len: 215 to MEMORY
2025-04-29 15:29:15,357 INFO reduce.EventFetcher: EventFetcher is interrupted.. Returning
2025-04-29 15:29:15,359 INFO mapred.LocalJobrunner: 1 / 1 copied.
2025-04-29 15:29:15,359 INFO reduce.MergeManagerImpl: finalMerge called with 1 in-memory map-outputs and 0 on-disk map-outputs
2025-04-29 15:29:15,361 INFO mapred.Merger: Merged 1 segments
2025-04-29 15:29:15,362 INFO mapred.Merger: Down to the last merge-pass, with 1 segments left of total size: 205 bytes
2025-04-29 15:29:15,362 INFO mapred.MergerManagerImpl: Merged 1 segments, 211 bytes to disk to satisfy reduce memory limit
2025-04-29 15:29:15,363 INFO mapred.Merger: Merging 1 files, 215 bytes from disk
2025-04-29 15:29:15,363 INFO mapred.Merger: Merging 0 segments, 0 bytes from memory into reduce
2025-04-29 15:29:15,363 INFO mapred.Merger: Merging 1 sorted segments
2025-04-29 15:29:15,365 INFO mapred.Merger: Down to the last merge-pass, with 1 segments left of total size: 205 bytes
2025-04-29 15:29:15,365 INFO mapred.LocalJobrunner: 1 / 1 copied.
2025-04-29 15:29:15,416 INFO mapred.Task: Task attempt_local537170479_0001_r_000000_0 is done. And is in the process of committing
2025-04-29 15:29:15,416 INFO mapred.Task: Task attempt_local537170479_0001_r_000000_0 is allowed to commit now
2025-04-29 15:29:15,423 INFO output.FileOutputCommitter: Saved output of task 'attempt_local537170479_0001_r_000000_0' to hdfs://localhost:9000/output
2025-04-29 15:29:15,424 INFO mapred.LocalJobrunner: reduce > reduce
2025-04-29 15:29:15,424 INFO mapred.Task: Task attempt_local537170479_0001_r_000000_0 done.
2025-04-29 15:29:15,424 INFO mapred.Task: Final Counters for attempt_local537170479_0001_r_000000_0: Counters: 30
File System Counters
FILE: Number of bytes read=4618
FILE: Number of bytes written=642686
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes written=69
HDFS: Number of read operations=10
HDFS: Number of write operations=3
HDFS: Number of bytes read erasure-coded=0
Map-Reduce Framework
Combine input records=0
Combine output records=0
Reduce input groups=10
Reduce shuffle bytes=215
Reduce input records=20
Reduce output records=10
Spilled Records=0
shuffled Maps=1
Failed Shuffles=0
Merged Map outputs=1

```

```

Activities Terminal Apr 29 15:33
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~

File System Counters
FILE: Number of bytes read=4618
FILE: Number of bytes written=422686
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=98
HDFS: Number of bytes written=69
HDFS: Number of read operations=10
HDFS: Number of large read operations=0
HDFS: Number of write operations=3
HDFS: Number of bytes read erasure-coded=0

Map-Reduce Framework
Combine input records=0
Combine output records=0
Group input records=10
Reduce shuffle bytes=215
Reduce input records=20
Reduce output records=10
Spilled Records=20
Shuffled Maps =1
Failed Shuffles=0
Merged Map outputs=1
GC time elapsed (ns)=0
Total committed heap usage (bytes)=633339904

Shuffle Errors
BAD_ID=0
CONNECTION=0
IO_ERROR=0
WRONG_LENGTH=0
WRONG_MAP=0
WRONG_REDUCE=0

File Output Format Counters
2025-04-29 15:29:15,424 INFO mapred.LocalJobRunner: Finishing task: attempt_local537170479_0001_r_000000_0
2025-04-29 15:29:15,424 INFO mapred.LocalJobRunner: reduce task executor complete.
2025-04-29 15:29:16,185 INFO mapreduce.Job: Job job_local537170479_0001 running in uber mode : false
2025-04-29 15:29:16,185 INFO mapreduce.Job: map 100% reduce 100%
2025-04-29 15:29:16,185 INFO mapreduce.Job: Job job_local537170479_0001 completed successfully
2025-04-29 15:29:16,185 INFO mapreduce.Job: Counters: 36
File System Counters
FILE: Number of bytes read=8774
FILE: Number of bytes written=1285157
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=108
HDFS: Number of bytes written=69
HDFS: Number of read operations=15
HDFS: Number of large read operations=0
HDFS: Number of write operations=4
HDFS: Number of bytes read erasure-coded=0

Map-Reduce Framework
Map input records=6
Map output records=20
Map output bytes=169

Activities Terminal Apr 29 15:34
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~

2025-04-29 15:29:15,424 INFO mapred.LocalJobRunner: reduce task executor complete.
2025-04-29 15:29:16,185 INFO mapreduce.Job: Job job_local537170479_0001 running in uber mode : false
2025-04-29 15:29:16,185 INFO mapreduce.Job: map 100% reduce 100%
2025-04-29 15:29:16,185 INFO mapreduce.Job: Job job_local537170479_0001 completed successfully
2025-04-29 15:29:16,185 INFO mapreduce.Job: Counters: 36
File System Counters
FILE: Number of bytes read=8774
FILE: Number of bytes written=1285157
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=180
HDFS: Number of bytes written=69
HDFS: Number of read operations=15
HDFS: Number of large read operations=0
HDFS: Number of write operations=4
HDFS: Number of bytes read erasure-coded=0

Map-Reduce Framework
Map input records=6
Map output records=20
Map output bytes=169
Map output compressed bytes=215
Input split bytes=86
Combine input records=0
Combine output records=0
Reduce input groups=10
Reduce shuffle bytes=215
Reduce input records=20
Reduce output Records=10
Spilled Records=40
Shuffled Maps =1
Failed Shuffles=0
Merged Map outputs=1
GC time elapsed (ns)=0
Total committed heap usage (bytes)=1266679808

Shuffle Errors
BAD_ID=0
CONNECTION=0
IO_ERROR=0
WRONG_LENGTH=0
WRONG_MAP=0
WRONG_REDUCE=0

File Input Format Counters
Bytes Read=90
File Output Format Counters
Bytes Written=0

0
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: $ hadoop fs -ls /output
Found 2 items
-rw-r--r-- 1 hadoop supergroup 0 2025-04-29 15:29 /output/_SUCCESS
-rw-r--r-- 1 hadoop supergroup 69 2025-04-29 15:29 /output/part-00000
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: $ hadoop fs -cat /output/part-00000
arun
brother 1
family 1
hi 1

```

```

Activities Terminal Apr 29 15:34
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
FILE: Number of bytes written=0
HDFS: Number of bytes written=69
HDFS: Number of read operations=15
HDFS: Number of large read operations=0
HDFS: Number of write operations=4
HDFS: Number of bytes read erasure-coded=0
Map-Reduce Framework
  Map output records=2
  Map output bytes=169
  Map output materialized bytes=215
  Input split bytes=80
  Combine output records=0
  Combine output bytes=0
  Reduce input groups=10
  Reduce shuffle bytes=215
  Reduce input records=20
  Reduce output records=10
  Spilled Records=40
  Shuffled Maps=0
  Failed Shuffles=0
  Merged Map outputs=1
  GC time elapsed (ns)=0
  Total committed heap usage (bytes)=12666679808
Shuffle Errors
  BAD_DNS=0
  CONNECTION=0
  IO_ERROR=0
  WRONG_LENGTH=0
  WRONG_MAP=0
  WRONG_REDUCE=0
File Input Format Counters
  Bytes Read=90
File Output Format Counters
  Bytes Written=69
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~$ hadoop fs -ls /output
Found 1 items
-rw-r--r--  1 hadoop supergroup          0 2025-04-29 15:29 /output/_SUCCESS
-rw-r--r--  1 hadoop supergroup         69 2025-04-29 15:29 /output/part-00000
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~$ hadoop fs -cat /output/part-00000
are 1
brother 1
empty 1
hi 1
how 5
ts 4
job 1
sister 1
you 1
your 4
hadoop@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~

```

The screenshot shows the Eclipse IDE interface with the following details:

- Title Bar:** eclipse-workspace - WordCount/src/MyDriver.java - Eclipse IDE
- Toolbar:** File Edit Source Refactor Navigate Search Project Run Window Help
- Package Explorer:** Shows the project structure for "WordCount".
- Code Editor:** Displays the Java code for `MyDriver.java`. The code implements a Driver class that extends `Configured` and implements `Tool`. It sets up a JobConf, configures input and output formats, and runs the job.
- Outline View:** Shows the outline of the `MyDriver` class.
- Task List:** Shows a single task: "Find All > Activate..."
- Problems View:** Shows a message: "Java 24 has been released. Would you like to install Eclipse Java Development Tools support for Java 24 into your IDE?" with options: "Install Java 24 Support", "Remind Me Later", and "Don't Ask Again".
- Bottom Status Bar:** Shows "Download Temurin JDK 24" and "Download OpenJDK 24".

```

1 import java.io.IOException;
2 import org.apache.hadoop.conf.Configured;
3 import org.apache.hadoop.fs.Path;
4 import org.apache.hadoop.mapreduce.Job;
5 import org.apache.hadoop.io.Text;
6 import org.apache.hadoop.mapred.FileInputFormat;
7 import org.apache.hadoop.mapred.FileOutputFormat;
8 import org.apache.hadoop.mapred.JobClient;
9 import org.apache.hadoop.mapred.Mapper;
10 import org.apache.hadoop.util.Tool;
11 import org.apache.hadoop.util.ToolRunner;
12
13 public class MyDriver extends Configured implements Tool {
14     public int run(String[] args) throws IOException {
15         if (args.length < 2) {
16             System.out.println("Please provide input and output paths");
17             return -1;
18         }
19         JobConf conf = new JobConf(MyDriver.class);
20         FileInputFormat.setInputPaths(conf, new Path(args[0]));
21         FileOutputFormat.setOutputPath(conf, new Path(args[1]));
22         conf.setMapperClass(MyMapper.class);
23         conf.setReducerClass(MyReducer.class);
24         conf.setMapOutputKeyClass(Text.class);
25         conf.setMapOutputValueClass(IntWritable.class);
26         conf.setOutputKeyClass(Text.class);
27         conf.setOutputValueClass(IntWritable.class);
28         JobClient.runJob(conf);
29         return 0;
30     }
31     public static void main(String[] args) throws Exception {
32         int exitCode = ToolRunner.run(new MyDriver(), args);
33         System.out.println(exitCode);
34     }
35 }
36

```

Activities Eclipse Apr 29 15:35

eclipse-workspace - WordCount/src/MyMapper.java - Eclipse IDE

File Edit Source Refactor Navigate Search Project Run Window Help

Package Explorer X JRE System Library [JavaSE-1.8] MyMapper.java MyDriver.java

src (default package) MyDriver.java MyMapper.java MyReducer.java

Referenced Libraries settings classpath .project

Task List X Find All Activate... ?

Outline X MyMapper

```
1 import java.io.IOException;
2 import org.apache.hadoop.io.IntWritable;
3 import org.apache.hadoop.io.LongWritable;
4 import org.apache.hadoop.io.Text;
5 import org.apache.hadoop.mapreduce.Mapper;
6 import org.apache.hadoop.mapred.OutputCollector;
7 import org.apache.hadoop.mapred.Reporter;
8
9
10 public class MyMapper extends MapReduceBase implements Mapper<LongWritable, Text, Text, IntWritable> {
11     public void map(LongWritable key, Text value, OutputCollector<Text, IntWritable> output, Reporter rep) throws IOException {
12         String line = value.toString();
13         for (String word : line.split(" ")) {
14             if (word.length() > 0) {
15                 output.collect(new Text(word), new IntWritable(1));
16             }
17         }
18     }
19 }
20
21 }
```

Problems Javadoc Declaration Install Java 24 Support X

Java 24 has been released.

Would you like to install Eclipse Java Development Tools support for Java 24 into your IDE?

Install Java 24 Support Remind Me Later Don't Ask Again

Would you like to download a Java 24 JDK?

Download Temurin JDK 24 Download OpenJDK 24

Writable Smart Insert 21:1:779

Activities Eclipse Apr 29 15:35

eclipse-workspace - WordCount/src/MyReducer.java - Eclipse IDE

File Edit Source Refactor Navigate Search Project Run Window Help

Package Explorer X JRE System Library [JavaSE-1.8] MyMapper.java MyReducer.java MyDriver.java

src (default package) MyDriver.java MyMapper.java MyReducer.java

Referenced Libraries settings classpath .project

Task List X Find All Activate... ?

Outline X MyReducer

```
1 import java.io.IOException;
2 import java.util.Iterator;
3 import org.apache.hadoop.io.IntWritable;
4 import org.apache.hadoop.io.Text;
5 import org.apache.hadoop.mapred.MapReduceBase;
6 import org.apache.hadoop.mapred.OutputCollector;
7 import org.apache.hadoop.mapred.Reducer;
8 import org.apache.hadoop.mapred.Reporter;
9
10 public class MyReducer extends MapReduceBase implements Reducer<Text, IntWritable, Text, IntWritable> {
11     public void reduce(Text key, Iterator<IntWritable> values, OutputCollector<Text, IntWritable> output, Reporter rep) throws IOException {
12         int count = 0;
13         while (values.hasNext()) {
14             count += values.next().get();
15         }
16         output.collect(key, new IntWritable(count));
17     }
18 }
```

Problems Javadoc Declaration Install Java 24 Support X

Java 24 has been released.

Would you like to install Eclipse Java Development Tools support for Java 24 into your IDE?

Install Java 24 Support Remind Me Later Don't Ask Again

Would you like to download a Java 24 JDK?

Download Temurin JDK 24 Download OpenJDK 24

Writable Smart Insert 19:1:726

Activities Eclipse Apr 29 15:35

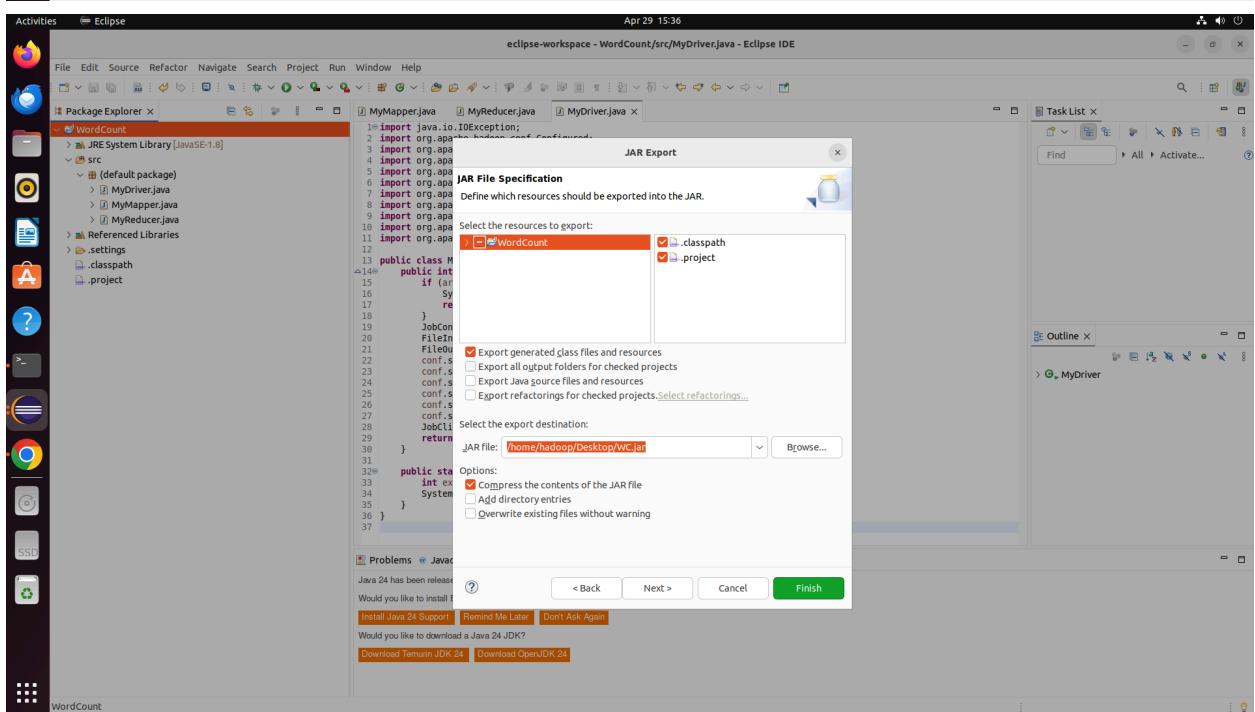
eclipse-workspace - WordCount/src/MyDriver.java - Eclipse IDE

```

File Edit Source Refactor Navigate Search Project Run Window Help
Package Explorer X eclipse-workspace - WordCount/src/MyDriver.java X Task List X
JRE System Library [JavaSE-1.8]
src (default package)
MyDriver.java
MyMapper.java
MyReducer.java
Referenced Libraries
settings
.classpath
.project
Find All Activate...
Outline X
MyDriver
Find All Activate...
Problems Javadoc Declaration Install Java 24 Support X
Java 24 has been released.
Would you like to install Eclipse Java Development Tools support for Java 24 into your IDE?
Install Java 24 Support Remind Me Later Don't Ask Again
Would you like to download a Java 24 JDK?
Download Temurin JDK 24 Download OpenJDK 24
Writable Smart Insert 37 : 1:1391 ...

```

1 import java.io.IOException;
2 import org.apache.hadoop.conf.Configuration;
3 import org.apache.hadoop.fs.Path;
4 import org.apache.hadoop.io.IntWritable;
5 import org.apache.hadoop.mapreduce.InputFormat;
6 import org.apache.hadoop.mapred.FileInputFormat;
7 import org.apache.hadoop.mapred.FileOutputFormat;
8 import org.apache.hadoop.mapred.JobClient;
9 import org.apache.hadoop.mapred.JobConf;
10 import org.apache.hadoop.util.Tool;
11 import org.apache.hadoop.util.ToolRunner;
12
13 public class MyDriver extends Configuration implements Tool {
14 public int run(String[] args) throws IOException {
15 if (args.length < 2) {
16 System.out.println("Please provide input and output paths");
17 return -1;
18 }
19 JobConf conf = new JobConf(MyDriver.class);
20 FileInputFormat.setInputPaths(conf, new Path(args[0]));
21 FileOutputFormat.setOutputPath(conf, new Path(args[1]));
22 conf.setMapperClass(MyMapper.class);
23 conf.setReducerClass(MyReducer.class);
24 conf.setOutputKeyClass(Text.class);
25 conf.setOutputValueClass(IntWritable.class);
26 conf.setMapOutputValueClass(IntWritable.class);
27 conf.setOutputKeyClass(Text.class);
28 conf.setOutputValueClass(IntWritable.class);
29 JobClient.runJob(conf);
30 }
31
32 public static void main(String[] args) throws Exception {
33 int exitCode = ToolRunner.run(new MyDriver(), args);
34 System.out.println(exitCode);
35 }
36 }



EXPERIMENT-7

From the following link extract the weather data

<https://github.com/tomwhite/hadoop-book/tree/master/input/ncdc/all>

- a) Create a MapReduce program to
find average temperature for each year from NCDC
data set.
- b) find the mean max temperature for every month.

Code with screenshots:

eclipse-workspace the - NCDCWeatherAnalysis/src/temp/AverageDriver.java - Eclipse IDE

```

File Edit Source Refactor Navigate Search Project Run Window Help
Package Explorer X
  NCDCWeatherAnalysis
    JRE System Library [JavaSE-1.8]
      src
        meanmax
          MeanMaxDriver.java
            AverageDriver.java
              AverageMapper.java
                AverageReducer.java
      temp
        AverageDriver.java
          AverageMapper.java
            AverageReducer.java
      Referenced Libraries
  Task List X
    Find All Activate...
  Connect Mylyn
    Connect to your task and ALM tools or create a local task.
  Outline X
    temp
      AverageDriver
        main(String[])
  Problems Javadoc Declaration Console X Install Java 24 Support
  <terminated> AverageDriver [Java Application] /usr/lib/jvm/java-8-openjdk-amd64/bin/java (06-May-2025, 2:49:49 pm - 2:49:51 pm elapsed: 0:00:01.788) [pid: 7013]
  Please enter input and output parameters
  ds
  ds
meanmax.MeanMaxDriver.java - NCDCWeatherAnalysis/src

```

eclipse-workspace the - NCDCWeatherAnalysis/src/temp/AverageDriver.java - Eclipse IDE

```

File Edit Source Refactor Navigate Search Project Run Window Help
Package Explorer X
  NCDCWeatherAnalysis
    JRE System Library [JavaSE-1.8]
      src
        meanmax
          MeanMaxDriver.java
            MeanMaxMapper.java
              AverageDriver.java
                AverageMapper.java
                  AverageReducer.java
      temp
        AverageDriver.java
          AverageMapper.java
            AverageReducer.java
      Referenced Libraries
  Task List X
    Find All Activate...
  Connect Mylyn
    Connect to your task and ALM tools or create a local task.
  Outline X
    temp
      AverageDriver
        main(String[])
  Problems Javadoc Declaration Console X Install Java 24 Support
  <terminated> AverageDriver [Java Application] /usr/lib/jvm/java-8-openjdk-amd64/bin/java (06-May-2025, 2:49:49 pm - 2:49:51 pm elapsed: 0:00:01.788) [pid: 7013]
  Please enter input and output parameters
  ds
  ds
meanmax.MeanMaxMapper.java - NCDCWeatherAnalysis/src

```

eclipse-workspace the - NCDCWeatherAnalysis/src/temp/AverageDriver.java - Eclipse IDE

```

File Edit Source Refactor Navigate Search Project Run Window Help
Package Explorer X
  NCDCWeatherAnalysis
    JRE System Library [JavaSE-1.8]
    src
      meanmax
        MeanMaxDriver.java
        MeanMaxMapper.java
        MeanMaxReducer.java
    temp
      AverageDriver.java
      AverageMapper.java
      AverageReducer.java
    Referenced Libraries
  Task List X
    Find All Activate...
  Connect Mylyn
    Connect to your task and ALM tools or create a local task.
  Outline X
    temp
      AverageDriver.java
        main(String[])
  Problems Javadoc Declaration Console X Install Java 24 Support
  <terminated> AverageDriver [Java Application] /usr/lib/jvm/java-8-openjdk-amd64/bin/java (06-May-2025, 2:49:49 pm - 2:49:51 pm elapsed: 0:00:01.788) [pid: 7013]
  Please enter input and output parameters
  d
  ds

```

meanmax.MeanMaxReducer.java - NCDCWeatherAnalysis/src

eclipse-workspace the - NCDCWeatherAnalysis/src/temp/AverageDriver.java - Eclipse IDE

```

File Edit Source Refactor Navigate Search Project Run Window Help
Package Explorer X
  NCDCWeatherAnalysis
    JRE System Library [JavaSE-1.8]
    src
      meanmax
        MeanMaxDriver.java
        MeanMaxMapper.java
        MeanMaxReducer.java
    temp
      AverageDriver.java
      AverageMapper.java
      AverageReducer.java
    Referenced Libraries
  Task List X
    Find All Activate...
  Connect Mylyn
    Connect to your task and ALM tools or create a local task.
  Outline X
    temp
      AverageDriver.java
        main(String[])
  Problems Javadoc Declaration Console X Install Java 24 Support
  <terminated> AverageDriver [Java Application] /usr/lib/jvm/java-8-openjdk-amd64/bin/java (06-May-2025, 2:49:49 pm - 2:49:51 pm elapsed: 0:00:01.788) [pid: 7013]
  Please enter input and output parameters
  d
  ds

```

temp.AverageDriver.java - NCDCWeatherAnalysis/src

Activities Eclipse May 6 15:22

eclipse-workspace the - NCDCWeatherAnalysis/src/temp/AverageDriver.java - Eclipse IDE

File Edit Source Refactor Navigate Search Project Run Window Help

Package Explorer X

- NCDCWeatherAnalysis
- JRE System Library [JavaSE-1.8]
- src
 - meanmax
 - MeanMaxDriver.java
 - MeanMaxMapper.java
 - MeanMaxReducer.java
 - temp
 - AverageDriver.java
 - AverageMapper.java
 - AverageReducer.java
- Referenced Libraries

```

1 package temp;
2 import org.apache.hadoop.fs.Path;
3 import org.apache.hadoop.io.IntWritable;
4 import org.apache.hadoop.io.Text;
5 import org.apache.hadoop.mapreduce.Job;
6 import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
7 import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;
8 public class AverageDriver {
9   public static void main(String[] args) throws Exception {
10     if (args.length != 2) {
11       System.err.println("Please enter input and output parameters");
12       System.exit(1);
13     }
14     Job job = Job.getInstance();
15     job.setJarByClass(AverageDriver.class); // Set the main class for the job
16     job.setJobName("Average Temperature by Year"); // Set a name for the job
17     FileInputFormat.addInputPath(job, new Path(args[0]));
18     FileOutputFormat.setOutputPath(job, new Path(args[1]));
19
20     // Set the Mapper and Reducer classes
21     job.setMapperClass(AverageMapper.class);
22     job.setReducerClass(AverageReducer.class);
23
24     // Set the output key and value types
25     job.setOutputKeyClass(Text.class);
26     job.setOutputValueClass(IntWritable.class);
27
28     // Submit the job and wait for completion
29     System.exit(job.waitForCompletion(true) ? 0 : 1);
30   }
31 }
32 
```

Problems Javadoc Declaration Console x Install Java 24 Support <terminated> AverageDriver [Java Application] /usr/lib/jvm/java-8-openjdk-amd64/bin/java (06-May-2025, 2:49:49 pm - 2:49:51 pm elapsed: 0:00:01.788) [pid: 7013]

Please enter input and output parameters

temp.AverageMapper.java - NCDCWeatherAnalysis/src

Activities Eclipse May 6 15:23

eclipse-workspace the - NCDCWeatherAnalysis/src/temp/AverageDriver.java - Eclipse IDE

File Edit Source Refactor Navigate Search Project Run Window Help

Package Explorer X

- NCDCWeatherAnalysis
- JRE System Library [JavaSE-1.8]
- src
 - meanmax
 - MeanMaxDriver.java
 - MeanMaxMapper.java
 - MeanMaxReducer.java
 - temp
 - AverageDriver.java
 - AverageMapper.java
 - AverageReducer.java
- Referenced Libraries

```

1 package temp;
2 import org.apache.hadoop.fs.Path;
3 import org.apache.hadoop.io.IntWritable;
4 import org.apache.hadoop.io.Text;
5 import org.apache.hadoop.mapreduce.Job;
6 import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
7 import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;
8 public class AverageDriver {
9   public static void main(String[] args) throws Exception {
10     if (args.length != 2) {
11       System.err.println("Please enter input and output parameters");
12       System.exit(1);
13     }
14     Job job = Job.getInstance();
15     job.setJarByClass(AverageDriver.class); // Set the main class for the job
16     job.setJobName("Average Temperature by Year"); // Set a name for the job
17     FileInputFormat.addInputPath(job, new Path(args[0]));
18     FileOutputFormat.setOutputPath(job, new Path(args[1]));
19
20     // Set the Mapper and Reducer classes
21     job.setMapperClass(AverageMapper.class);
22     job.setReducerClass(AverageReducer.class);
23
24     // Set the output key and value types
25     job.setOutputKeyClass(Text.class);
26     job.setOutputValueClass(IntWritable.class);
27
28     // Submit the job and wait for completion
29     System.exit(job.waitForCompletion(true) ? 0 : 1);
30   }
31 }
32 
```

Problems Javadoc Declaration Console x Install Java 24 Support <terminated> AverageDriver [Java Application] /usr/lib/jvm/java-8-openjdk-amd64/bin/java (06-May-2025, 2:49:49 pm - 2:49:51 pm elapsed: 0:00:01.788) [pid: 7013]

Please enter input and output parameters

temp.AverageReducer.java - NCDCWeatherAnalysis/src

Activities Terminal May 6 15:25

```

hadoop@bmsccse-HP-Elite-Tower-600-G9-Desktop-PC: ~/Desktop
hadoop@bmsccse-HP-Elite-Tower-600-G9-Desktop-PC: $ hdfs dfs -mkdir /ncdc_input
mkdir: Call From bmsccse-HP-Elite-Tower-600-G9-Desktop-PC@127.0.1.1 to localhost:9000 failed on connection exception: java.net.ConnectException: Connection refused; For more details see: http://wiki.apache.org/hadoop/ConnectionRefused
hadoop@bmsccse-HP-Elite-Tower-600-G9-Desktop-PC: $ start-dfs.sh
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [bmsccse-HP-Elite-Tower-600-G9-Desktop-PC]
Starting resourcemanager
Starting nodemanagers
hadoop@bmsccse-HP-Elite-Tower-600-G9-Desktop-PC: $ jps
15742 SecondaryNameNode
15745 DataNode
15556 NameNode
7093 org.eclipse.equinox.launcher_1.6.1000.v20250227-1734.jar
16504 NodeManager
16937 Jps
16938 ResourceManager
hadoop@bmsccse-HP-Elite-Tower-600-G9-Desktop-PC: $ hdfs dfs -mkdir /ncdc_input
hadoop@bmsccse-HP-Elite-Tower-600-G9-Desktop-PC: $ git clone https://github.com/tomwhite/hadoop-book.git
fatal: destination path 'hadoop-book' already exists and is not an empty directory.
hadoop@bmsccse-HP-Elite-Tower-600-G9-Desktop-PC: $ cd hadoop-book/input/ncdc/all/
hdfs dfs -put * /ncdc_input
Found 2 items
-rw-r--r-- 1 hadoop supergroup 73867 2025-05-06 15:11 /ncdc_input/1901.gz
-rw-r--r-- 1 hadoop supergroup 74105 2025-05-06 15:11 /ncdc_input/1902.gz
hadoop@bmsccse-HP-Elite-Tower-600-G9-Desktop-PC: $ hadoop jar AverageTemp.jar temp.AverageDriver /ncdc_input /output_avg
JAR does not exist or is not a normal file: /home/bmsccse/hadoop/hadoop-book/input/ncdc/all/AverageTemp.jar
hadoop@bmsccse-HP-Elite-Tower-600-G9-Desktop-PC: $ hadoop jar AverageTemp.jar temp.AverageDriver /ncdc_input /output_avg
JAR does not exist or is not a normal file: /home/bmsccse/hadoop/Desktop/AverageTemp.jar
hadoop@bmsccse-HP-Elite-Tower-600-G9-Desktop-PC: $ hadoop jar NDCWeather.jar temp.AverageDriver /ncdc_input /output_avg
ACFr0g1t880DXZ9jzYb1lxQ8_J-Myt4QJozqp_tewZ05d059e0h@Eqf_um0I_d7nxM5b849Nx4Mg2fKz5Gak8WTKMyxCYNBKewr7y0GFllI7dBJC8Kkdjx9GAsRZJrJpzGQxTrpgSUCU.pdf
file1.txt
NDCWeather.jar
sample.txt
Tejaswini.jar
WordCount170.jar
WordCount.jar
hadoop@bmsccse-HP-Elite-Tower-600-G9-Desktop-PC: $ hadoop jar NDCWeather.jar temp.AverageDriver /ncdc_input /output_avg
2025-05-06 15:14:57,391 INFO impl.MetricsConfig: Loaded properties from hadoop-metrics2.properties
2025-05-06 15:14:57,428 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).
2025-05-06 15:14:57,429 INFO impl.MetricsSystemImpl: Jobtracker metrics system started
2025-05-06 15:14:57,485 WARN mapred.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
2025-05-06 15:14:57,533 INFO mapred.JobSubmitter: number of splits?
2025-05-06 15:14:57,607 INFO mapred.JobSubmitter: Submitting tokens for job: job_local1944593155_0001
2025-05-06 15:14:57,656 INFO mapred.JobSubmitter: Executing with tokens: []
2025-05-06 15:14:57,657 INFO mapred.JobSubmitter: The url to track the job: http://localhost:8088/
2025-05-06 15:14:57,658 INFO mapred.JobSubmitter: Job: Running job: job_local1944593155_0001
2025-05-06 15:14:57,661 INFO mapred.LocalJobRunner: OutputCommitter set in config null
2025-05-06 15:14:57,661 INFO mapred.LocalJobRunner: OutputCommitter: FileOutputCommitter Algorithm version is 2
2025-05-06 15:14:57,661 INFO mapred.LocalJobRunner: OutputCommitter: FileOutputCommitter skip cleanup temporary folders under output directory:false, ignore cleanup failures: false
2025-05-06 15:14:57,661 INFO mapred.LocalJobRunner: OutputCommitter: org.apache.hadoop.mapred.lib.output.FileOutputCommitter

```

Activities Terminal May 6 15:26

```

hadoop@bmsccse-HP-Elite-Tower-600-G9-Desktop-PC: ~/Desktop
hadoop@bmsccse-HP-Elite-Tower-600-G9-Desktop-PC: $ hadoop jar NDCWeather.jar temp.AverageDriver /ncdc_input /output_avg
2025-05-06 15:14:57,391 INFO impl.MetricsConfig: Loaded properties from hadoop-metrics2.properties
2025-05-06 15:14:57,428 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).
2025-05-06 15:14:57,439 INFO impl.MetricsSystemImpl: Jobtracker metrics system started
2025-05-06 15:14:57,451 INFO mapred.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
2025-05-06 15:14:57,537 INFO input.FileInputFormat: Total input files to process : 2
2025-05-06 15:14:57,601 INFO mapred.JobSubmitter: Submitting tokens for job: job_local1944593155_0001
2025-05-06 15:14:57,607 INFO mapred.JobSubmitter: Executing with tokens: []
2025-05-06 15:14:57,656 INFO mapred.JobSubmitter: The url to track the job: http://localhost:8088/
2025-05-06 15:14:57,657 INFO mapred.JobSubmitter: Job: Running job: job_local1944593155_0001
2025-05-06 15:14:57,661 INFO mapred.LocalJobRunner: OutputCommitter set in config null
2025-05-06 15:14:57,661 INFO mapred.LocalJobRunner: OutputCommitter: FileOutputCommitter Algorithm version is 2
2025-05-06 15:14:57,661 INFO mapred.LocalJobRunner: OutputCommitter: FileOutputCommitter skip cleanup temporary folders under output directory:false, ignore cleanup failures: false
2025-05-06 15:14:57,661 INFO mapred.LocalJobRunner: OutputCommitter: org.apache.hadoop.mapred.lib.output.FileOutputCommitter
2025-05-06 15:14:57,692 INFO mapred.LocalJobRunner: Waiting for map tasks
2025-05-06 15:14:57,701 INFO output.FileOutputCommitter: Starting task: attempt_local1944593155_00001_m_000000_0
2025-05-06 15:14:57,701 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2025-05-06 15:14:57,709 INFO mapred.Task: Using ResourceCalculatorProcessTree : []
2025-05-06 15:14:57,744 INFO mapred.Task: TaskAttemptID:attempt_local1944593155_00001_m_000000_0
2025-05-06 15:14:57,739 INFO mapred.MapTask: (EQUATOR) @ kvl_Z6214396(10485784)
2025-05-06 15:14:57,739 INFO mapred.MapTask: InputSplit bytes=104857600
2025-05-06 15:14:57,739 INFO mapred.MapTask: MapTask
2025-05-06 15:14:57,739 INFO mapred.MapTask: mapreduce.task.io.sort.mb: 100
2025-05-06 15:14:57,739 INFO mapred.MapTask: soft limit at 83860800
2025-05-06 15:14:57,739 INFO mapred.MapTask: bufstart = 0; bufvoid = 104857600
2025-05-06 15:14:57,739 INFO mapred.MapTask: kvstart = 26214396; length = 6553600
2025-05-06 15:14:57,739 INFO mapred.MapTask: mapreduce.map.output.collector.class = org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2025-05-06 15:14:57,754 INFO zlib.ZlibFactory: Successfully loaded & initialized native-zlib library
2025-05-06 15:14:57,754 INFO compress.CodecPool: Got brand-new decompressor [-g]
2025-05-06 15:14:57,815 INFO mapred.LocalJobRunner:
2025-05-06 15:14:57,816 INFO mapred.MapTask: Starting flush of map output
2025-05-06 15:14:57,816 INFO mapred.MapTask: Spilling map output
2025-05-06 15:14:57,816 INFO mapred.MapTask: mapEnd = 59085; bufvoid = 104857600
2025-05-06 15:14:57,816 INFO mapred.MapTask: kvstart = 26214396(10485784); kvend = 26180140(104752560); length = 20257/6553600
2025-05-06 15:14:57,822 INFO mapred.MapTask: finished spill 0
2025-05-06 15:14:57,826 INFO mapred.Task: Task:attempt_local1944593155_00001_m_000000_0 is done. And is in the process of committing
2025-05-06 15:14:57,828 INFO mapred.LocalJobRunner: map
2025-05-06 15:14:57,829 INFO mapred.Task: Task:attempt_local1944593155_00001_m_000000_0 is done.
2025-05-06 15:14:57,830 INFO mapred.Task: Final Counters for attempt_local1944593155_00001_m_000000_0: Counters: 23
File System Counters
FILE: Number of bytes read=7873
FILE: Number of bytes written=719826
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=74105
HDFS: Number of bytes written=0
HDFS: Number of read operations=5
HDFS: Number of large read operations=0
HDFS: Number of write operations=1
HDFS: Number of bytes read erasure-coded=0
Map-Reduce Framework
Map input records=6565
Map output records=6565
Map output bytes=59085

```

```
Activities Terminal May 16 15:26 
hadoop@bmsece-HP-Elite-Tower-600-G9-Desktop-PC: ~/Desktop

Input split bytes=105
  Block Input records=0
  Spilled Records=6564
  Failed Shuffles=0
  Merged Map Outputs=0
  Output cleanup bytes=0
  Total committed heap usage (bytes)=526385152
  File Input Format Counters
    Bytes Read=73867
2025-05-06 15:14:57.895 INFO mapred.LocalJobRunner: Finishing task: attempt_local1944593155_0001_m_000001_0
2025-05-06 15:14:57.895 INFO mapred.LocalJobRunner: map task executor complete.
2025-05-06 15:14:57.895 INFO mapred.LocalJobRunner: Starting commit for reduce tasks
2025-05-06 15:14:57.897 INFO mapred.LocalJobRunner: Committing for reduce tasks
2025-05-06 15:14:57.904 INFO output.FileOutputCommitter: File output Committer Algorithm version is 2
2025-05-06 15:14:57.904 INFO output.FileOutputCommitter: skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
2025-05-06 15:14:57.904 INFO mapred.Task: Using ResourceCalculatorProcessTree : []
2025-05-06 15:14:57.906 WARN org.apache.hadoop.metrics.MetricsSystemImpl: JobTracker metrics system already initialized
2025-05-06 15:14:57.911 INFO mapred.Merger: Merging 2 sorted segments, 2206 bytes from map-output for attempt_local1944593155_0001_r_000000_0
2025-05-06 15:14:57.916 INFO reduce.EventFetcher: localfetcher#0 Thread started: EventFetcher for fetching Map Completion Events
2025-05-06 15:14:57.927 INFO reduce.LocalFetcher: localfetcher#1 about to shuffle output of map attempt_local1944593155_0001_r_000001_0 decomp: 72206 len: 72210 to MEMORY
2025-05-06 15:14:57.928 INFO reduce.InMemoryMapOutputs: Read 72206 bytes from map-output for attempt_local1944593155_0001_r_000000_0
2025-05-06 15:14:57.929 INFO reduce.MergeManagerImpl: closeInMemoryFile -> map-output of size: 72206, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->72206
2025-05-06 15:14:57.930 INFO reduce.LocalFetcher: localfetcher#1 about to shuffle output of map attempt_local1944593155_0001_r_000000_0 decomp: 72217 len: 72221 to MEMORY
2025-05-06 15:14:57.930 INFO reduce.InMemoryMapOutputs: Read 72217 bytes from map-output for attempt_local1944593155_0001_m_000000_0
2025-05-06 15:14:57.930 INFO reduce.MergeManagerImpl: closeInMemoryFile -> map-output of size: 72217, inMemoryMapOutputs.size() -> 2, commitMemory -> 72206, usedMemory ->144423
2025-05-06 15:14:57.930 INFO reduce.EventFetcher: EventFetcher is interrupted. Returning
2025-05-06 15:14:57.930 WARN LocalJobRunner: Failed readhead on file
EBADF: Bad file descriptor
  at org.apache.hadoop.io.nativeio.NativeIO$PosIX.posix_fadvise(Native Method)
  at org.apache.hadoop.io.nativeio.NativeIO$PosIX.posix_fadviseIfPossible(NativeIO.java:419)
  at org.apache.hadoop.io.nativeio.NativeIO$PosIX$CacheManipulator.posix_fadviseIfPossible(NativeIO.java:296)
  at org.apache.hadoop.io.nativeio.NativeIO$PosIX$CacheManipulator.run(NativeIO$PosIX$CacheManipulator.java:101)
  at java.base/java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:1128)
  at java.base/java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:628)
  at java.base/java.lang.Thread.run(Thread.java:829)
2025-05-06 15:14:57.930 INFO mapred.LocalJobRunner: 2 / 2 copied.
2025-05-06 15:14:57.930 INFO reduce.MergeManagerImpl: finalMerge with 2 in-memory map-outputs and 0 on-disk map-outputs
2025-05-06 15:14:57.930 INFO mapred.Merger: Merging 2 sorted segments
2025-05-06 15:14:57.933 INFO mapred.Merger: Down to the last merge-pass, with 2 segments left of total size: 144409 bytes
2025-05-06 15:14:57.938 INFO reduce.MergeManagerImpl: Merged 2 segments, 144423 bytes to disk to satisfy reduce memory limit
2025-05-06 15:14:57.938 INFO reduce.MergeManagerImpl: Merging 1 files, 144425 bytes from disk
2025-05-06 15:14:57.938 INFO reduce.MergeManagerImpl: Merging 0 segments, 0 bytes from memory into reduce
2025-05-06 15:14:57.938 INFO mapred.Merger: Merging 1 sorted segments
2025-05-06 15:14:57.939 INFO mapred.Merger: Down to the last merge-pass, with 1 segments left of total size: 144414 bytes
2025-05-06 15:14:57.939 INFO mapred.LocalJobRunner: 2 / 2 copied.
2025-05-06 15:14:57.939 INFO mapred.JobInProgress: mapred.mapoutput.on is deprecated. Instead, use mapreduce.job.skiprecords
2025-05-06 15:14:58.022 INFO mapred.Task: Task attempt_local1944593155_0001_r_000000_0 is done. And its in the process of committing
2025-05-06 15:14:58.023 INFO mapred.LocalJobRunner: 2 / 2 copied.
2025-05-06 15:14:58.023 INFO mapred.Task: Task attempt_local1944593155_0001_r_000000_0 is allowed to commit now
2025-05-06 15:14:58.030 INFO mapred.FileOutputCommitter: Saved output of task 'attempt_local1944593155_0001_r_000000_0' to hdfs://localhost:9006/output_avg
2025-05-06 15:14:58.030 INFO mapred.LocalJobRunner: reduce > reduce
2025-05-06 15:14:58.030 INFO mapred.Task: Task attempt_local1944593155_0001_r_000000_0 done.
2025-05-06 15:14:58.030 INFO mapred.Task: Final Counters for attempt_local1944593155_0001_r_000000_0: Counters: 30
  File System Counters
    ...EJ/F. Number of bytes read=29702?
```

```

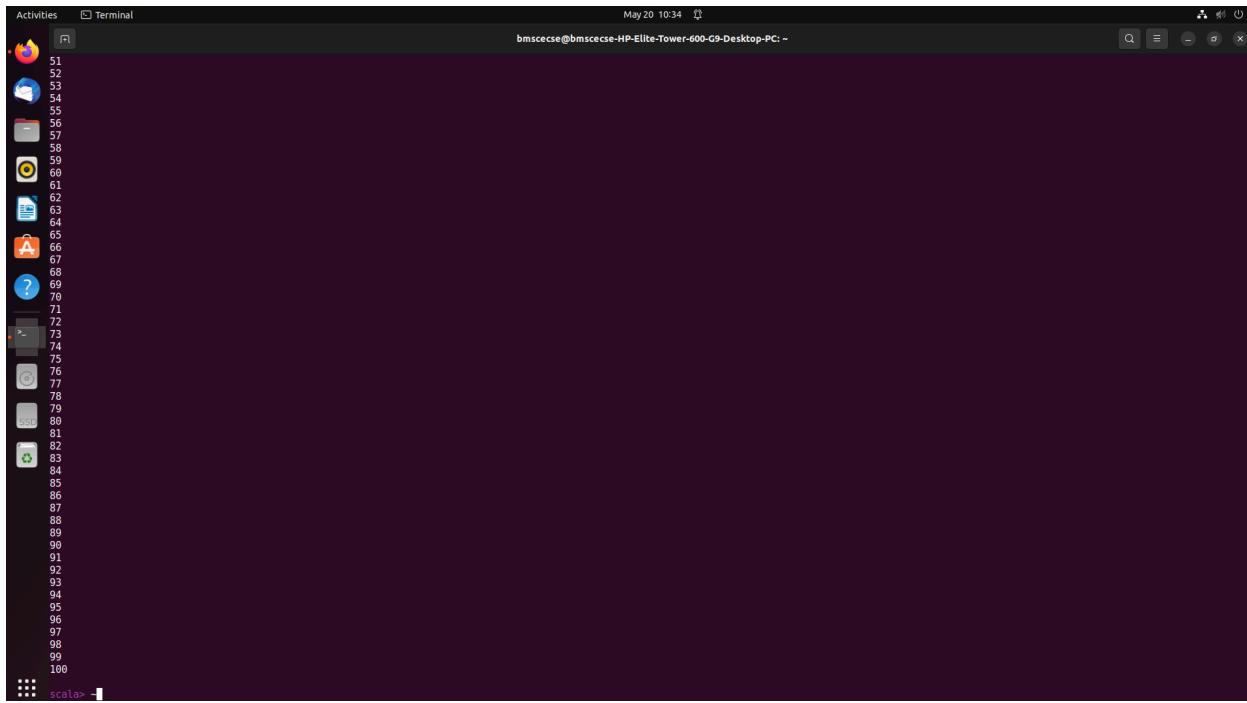
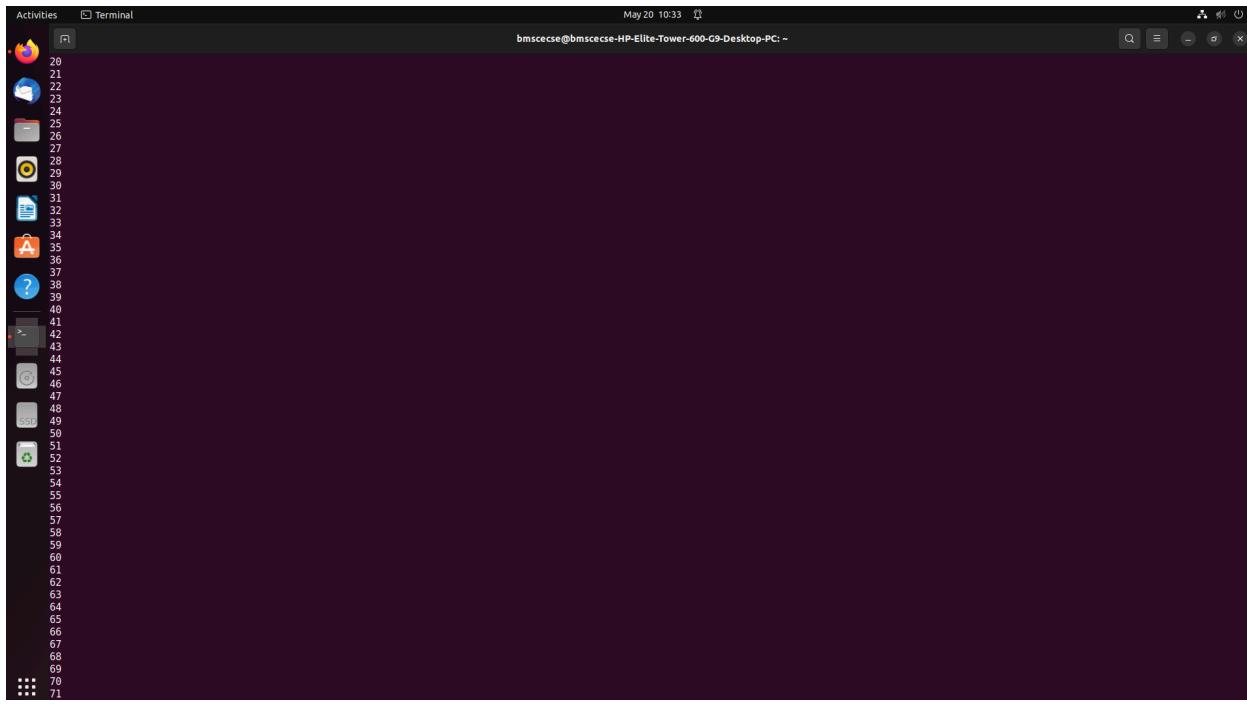
Activities Terminal May 6 15:27 hadoop@bmscsece-HP-Elite-Tower-600-G9-Desktop-PC: ~/Desktop
2025-05-06 15:14:58,038 INFO mapred.Task: Final Counters for attempt_local1944593155_0001_r_000000_0: Counters: 30
  FILE: Number of bytes read=297022
  FILE: Number of bytes written=936493
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=147972
  HDFS: Number of bytes written=16
  HDFS: Number of read operations=12
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=3
  HDFS: Number of bytes read erasure-coded=0
Map-Reduce Framework
  Combine Input records=0
  Combine output records=0
  Reduce Input groups=2
  Map Input records=144431
  Reduce Input records=13129
  Reduce output records=2
  Spilled Records=13129
  Shuffled Maps =2
  Failed Shuffles=0
  Merged Map outputs=2
  GC time elapsed (ms)=2
  Total committed heap usage (bytes)=633339904
Shuffle Errors
  BAD_ID=0
  CONNECTION=0
  IO_ERROR=0
  WRONG_LENGTH=0
  WRONG_MAP=0
  WRONG_REDUCE=0
File Output Format Counters
  Bytes Written=1
2025-05-06 15:14:58,038 INFO mapred.LocalJobRunner: Finishing task: attempt_local1944593155_0001_r_000000_0
2025-05-06 15:14:58,038 INFO mapred.LocalJobRunner: reduce task executor complete
2025-05-06 15:14:58,659 INFO mapreduce.Job: Job job_local1944593155_0001 running in uber mode : false
2025-05-06 15:14:58,659 INFO mapreduce.Job: map 100% reduce 100%
2025-05-06 15:14:58,664 INFO mapreduce.Job: Job job_local1944593155_0001 completed successfully
2025-05-06 15:14:58,664 INFO mapreduce.Job: Counters: 36
  File System Counters
    FILE: Number of bytes read=312997
    FILE: Number of bytes written=2448387
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=1000000
    HDFS: Number of bytes written=16
    HDFS: Number of read operations=24
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=5
    HDFS: Number of bytes read erasure-coded=0
  Map-Reduce Framework
    Map output Records=13129
    Map output records=13129
May 6 15:27 hadoop@bmscsece-HP-Elite-Tower-600-G9-Desktop-PC: ~/Desktop
Reduce output records=12
Spilled Records=26258
Shuffled Maps =2
Failed Shuffles=0
Merged Map outputs=2
GC time elapsed (ms)=3
Total committed heap usage (bytes)=1686110208
Shuffle Errors
  BAD_ID=0
  CONNECTION=0
  IO_ERROR=0
  WRONG_LENGTH=0
  WRONG_MAP=0
  WRONG_REDUCE=0
File Input Format Counters
  Bytes Read=147972
File Output Format Counters
  Bytes Written=72
hadoop@bmscsece-HP-Elite-Tower-600-G9-Desktop-PC: ~/Desktop$ hdfs dfs -ls /output_avg
Found 2 items
-rw-r--r--  1 hadoop supergroup      0 2025-05-06 15:14 /output_avg/_SUCCESS
-rw-r--r--  1 hadoop supergroup     16 2025-05-06 15:14 /output_avg/part-r-00000
hadoop@bmscsece-HP-Elite-Tower-600-G9-Desktop-PC: ~/Desktop$ hdfs dfs -cat /output_avg/part-r-00000
1901 46
1902 21
hadoop@bmscsece-HP-Elite-Tower-600-G9-Desktop-PC: ~/Desktop$ hdfs dfs -ls /output_avg
Found 2 items
-rw-r--r--  1 hadoop supergroup      0 2025-05-06 15:14 /output_avg/_SUCCESS
-rw-r--r--  1 hadoop supergroup     16 2025-05-06 15:14 /output_avg/part-r-00000
hadoop@bmscsece-HP-Elite-Tower-600-G9-Desktop-PC: ~/Desktop$ hdfs dfs -cat /output_avg/part-r-00000
1901 46
1902 21
hadoop@bmscsece-HP-Elite-Tower-600-G9-Desktop-PC: ~/Desktop$ hdfs dfs -cat /output_mean/part-r-00000
01 4
02 1
03 6
04 34
05 89
06 143
07 182
08 172
09 123
10 73
11 21
12 3
hadoop@bmscsece-HP-Elite-Tower-600-G9-Desktop-PC: ~/Desktop$ hadoop@bmscsece-HP-Elite-Tower-600-G9-Desktop-PC: ~/Desktop$ hdfs dfs -cat /output_avg/part-r-00000
1901 46
1902 21
hadoop@bmscsece-HP-Elite-Tower-600-G9-Desktop-PC: ~/Desktop$ hdfs dfs -cat /output_avg/part-r-00000
1901 46
1902 21

```

EXPERIMENT-8

Write a Scala program to print numbers from 1 to 100 using for loop.

Code with screenshots:



EXPERIMENT-9

Using RDD and FlatMap count how many times each word appears in a file and write out a list of words whose count is strictly greater than 4 using Spark.

Code with Screenshots:

```
bmscecse@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ spark-shell
25/05/20 15:32:38 WARN Utils: Your hostname, bmscecse-HP-Elite-Tower-800-G9-Desktop-PC resolves to a loopback address: 127.0.1.1
; using 10.124.2.8 instead (on interface eno1)
25/05/20 15:32:38 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.spark.unsafe.Platform (file:/opt/spark/jars/spark-unsafe_2.12-3.0.3.jar) to constructor java.nio.DirectByteBuffer(long,int)
WARNING: Please consider reporting this to the maintainers of org.apache.spark.unsafe.Platform
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
25/05/20 15:32:38 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
Spark context Web UI available at http://10.124.2.8:4040
Spark context available as 'sc' (master = local[*], app id = local-1747735361481).
Spark session available as 'spark'.
Welcome to

    ____/
   / \  - \ \  - \ / \ / \
  /_ \ / . \ \ / / \ \ \ \
             version 3.0.3
             /_ \

Using Scala version 2.12.10 (OpenJDK 64-Bit Server VM, Java 11.0.26)
Type in expressions to have them evaluated.
Type :help for more information.

scala> val textFile = sc.textFile("/home/bmscecse/Desktop/sparkdata.txt")
textFile: org.apache.spark.rdd.RDD[String] = /home/bmscecse/Desktop/sparkdata.txt MapPartitionsRDD[1] at textFile at <console>:24
4

scala>

scala> val counts = textFile
counts: org.apache.spark.rdd.RDD[String] = /home/bmscecse/Desktop/sparkdata.txt MapPartitionsRDD[1] at textFile at <console>:24

scala>   .flatMap(line => line.split(" "))
res0: org.apache.spark.rdd.RDD[String] = MapPartitionsRDD[2] at flatMap at <console>:26

scala>   .map(word => (word, 1))

scala> val data = sc.textFile("sparkdata.txt")
data: org.apache.spark.rdd.RDD[String] = sparkdata.txt MapPartitionsRDD[1] at textFile at <console>:25

scala> val splitdata = data.flatMap(line => line.split(" "))
splitdata: org.apache.spark.rdd.RDD[String] = MapPartitionsRDD[2] at flatMap at <console>:26

scala> val mapdata = splitdata.map(word => (word, 1))
mapdata: org.apache.spark.rdd.RDD[(String, Int)] = MapPartitionsRDD[3] at map at <console>:26

scala> val reducedata = mapdata.reduceByKey(_ + _)
reducedata: org.apache.spark.rdd.RDD[(String, Int)] = ShuffledRDD[4] at reduceByKey at <console>:26

scala> reducedata.collect.foreach(println)
(,1)
(hello,2)
(world,1)
(spark,1)
```

EXPERIMENT-10

Write a simple streaming program in Spark to receive text data streams on a particular port, perform basic text cleaning (like white space removal, stop words removal, lemmatization, etc.), and print the cleaned text on the screen. (Open Ended Question).

Code with Screenshots:

```
import org.apache.spark.SparkConf  
  
import org.apache.spark.streaming.{Seconds, StreamingContext}  
  
import org.apache.spark.streaming.dstream.DStream  
  
  
object StreamingTextCleaner {  
  
    val stopWords = Set(  
        "a", "an", "the", "is", "are", "was", "were", "this", "that",  
        "on", "in", "and", "or", "of", "to", "for", "with", "as", "by"  
    )  
  
  
    def simpleLemmatize(word: String): String = {  
        if (word.endsWith("ing")) word.dropRight(3)  
        else if (word.endsWith("ed")) word.dropRight(2)  
        else word  
    }  
}
```

```

def main(args: Array[String]): Unit = {
  if (args.length < 2) {
    System.err.println("Usage: StreamingTextCleaner <hostname> <port>")
    System.exit(1)
  }

  val hostname = args(0)
  val port = args(1).toInt

  val conf = new SparkConf().setAppName("StreamingTextCleaner").setMaster("local[*]")
  val ssc = new StreamingContext(conf, Seconds(5))

  val lines = ssc.socketTextStream(hostname, port)

  val cleanedLines: DStream[String] = lines.map(_.toLowerCase)
    .map(_.trim)
    .map(line => line.replaceAll("[^a-zA-Z\\s]", ""))
    .flatMap(_.split("\\s+"))
    .filter(word => word.nonEmpty && !stopWords.contains(word))
    .map(simpleLemmatize)
    .filter(_.nonEmpty)
    .reduceByWindow(

```

```

(a: String, b: String) => a + " " + b,
Seconds(5),
Seconds(5)

)

cleanedLines.print()

ssc.start()
ssc.awaitTermination()
}

}

Requirement already satisfied: nltk in /usr/local/lib/python3.11/dist-packages (3.9.1)
Requirement already satisfied: click in /usr/local/lib/python3.11/dist-packages (from nltk) (8.2.0)
Requirement already satisfied: joblib in /usr/local/lib/python3.11/dist-packages (from nltk) (1.5.0)
Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.11/dist-packages (from nltk) (2024.11.6)
Requirement already satisfied: tqdm in /usr/local/lib/python3.11/dist-packages (from nltk) (4.67.1)
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data]   Unzipping tokenizers/punkt.zip.
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Unzipping corpora/stopwords.zip.
[nltk_data] Downloading package wordnet to /root/nltk_data...
+-----+
|word |
+-----+
|hello |
|hate |
|hate |
|love |
|dont |
|want |
|cant |
|put |
|nobody|
|else |
+-----+

```