

1BM22CS151

8a)

INPUT:

```
GNU nano 6.2
10:00,abc,P1,300ms
12:00,abac,P1,100ms
12:10,azxc,P2,6000ms
14:00,aabc,P1,1100ms
15:00,acbc,P1,1100ms
15:10,aznxc,P2,6000ms
16:00,azxc,P3,6000ms
16:10,azdxc,P3,100000ms
17:00,azaaxc,P4,6000ms
17:10,azadxc,P6,100000ms
18:00,azagtaxc,P4,6000ms
```

CODES:

```
hadoop@bmscece-HP-Elite-Tower-800-G9-Desktop-PC: $ sudo apt update
sudo apt install python3 python3-pip -y
pip3 install mrjob
[sudo] password for hadoop:
Hit:2 http://in.archive.ubuntu.com/ubuntu jammy InRelease
Hit:3 http://security.ubuntu.com/ubuntu jammy-security InRelease
Hit:4 http://in.archive.ubuntu.com/ubuntu jammy-updates InRelease
Hit:5 https://repo.mongodb.org/apt/ubuntu jammy/mongodb-org/7.0 InRelease
Hit:6 http://in.archive.ubuntu.com/ubuntu jammy-backports InRelease
Ign:1 https://downloads.apache.org/cassandra/debian 40x InRelease
Err:8 https://downloads.apache.org/cassandra/debian 40x Release
      40x Not Found [IP: 135.181.214.104 443]
Hit:7 https://apache.jfrog.io/artifactory/cassandra-deb 41x InRelease
Reading package lists... Done
E: The repository 'http://www.apache.org/dist/cassandra/debian 40x Release' does not have a Release file.
N: Updating from such a repository can't be done securely, and is therefore disabled by default.
N: See apt-secure(8) manpage for repository creation and user configuration details.
M: https://deb.debian.cassandra.apache.org/dist/41x InRelease: Key is stored in legacy trusted.gpg keyring (/etc/apt/trusted.gpg), see the DEPRECATION section in apt-key(8) for details.
M: Target Packages (main/binary-and64/Packages) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target Packages (main/binary-i386/Packages) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target Packages (main/binary-all/Packages) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target Translations (main/i18n/Translation-en-IN) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target Translations (main/i18n/Translation-en) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target DEP-11 (main/dep11/Components-and64.yml) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target DEP-11 (main/dep11/Components-all.yml) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target DEP-11-icons-small (main/dep11/icons-48x48.tar) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target DEP-11-icons (main/dep11/icons-64x64.tar) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target DEP-11-icons-hidpi (main/dep11/icons-64x64@2.tar) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target CNF (main/cnf/Commands-and64) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target CNF (main/cnf/Commands-all) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target Packages (main/binary-and64/Packages) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target Packages (main/binary-i386/Packages) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target Packages (main/binary-all/Packages) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target Translations (main/i18n/Translation-en-IN) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target Translations (main/i18n/Translation-en) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target DEP-11 (main/dep11/Components-and64.yml) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target DEP-11 (main/dep11/Components-all.yml) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target DEP-11-icons-small (main/dep11/icons-48x48.tar) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target DEP-11-icons (main/dep11/icons-64x64.tar) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target DEP-11-icons-hidpi (main/dep11/icons-64x64@2.tar) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target CNF (main/cnf/Commands-and64) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
M: Target CNF (main/cnf/Commands-all) is configured multiple times in /etc/apt/sources.list.d/cassandra.sources.list:1 and /etc/apt/sources.list.d/cassandra.sources.list:2
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
python3 is already the newest version (3.10.6-1-22.04.1).
python3 set to manually installed.
python3-pip is already the newest version (22.0.2+dfsg-1ubuntu0.5).
The following packages were automatically installed and are no longer required:
  libwp6-1.0-1 libwp6backend-fdo-1.0-1
Use 'sudo apt autoremove' to remove them.
0 upgraded, 0 newly installed, 0 to remove and 19 not upgraded.
Defaulting to user installation because normal site-packages is not writeable
Collecting mrjob
  Downloading mrjob-0.7.4-py2.py3-none-any.whl (439 kB)
    430.6/439.6 KB 3.0 MB/s eta 0:00:00
```

```

hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ nano ~/pagecount_mrjob.py
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ $ nano ~/pagecount_mrjob.py
$: command not found
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ nano ~/pagecount_mrjob.py
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ python3 ~/pagecount_mrjob.py ~/input.txt
No configs found; falling back on auto-configuration
No configs specified for inline runner
Creating temp directory /tmp/pagecount_mrjob.hadoop.20250527.094917.716969
Running step 1 of 1...
job output is in /tmp/pagecount_mrjob.hadoop.20250527.094917.716969/output
Streaming final output from /tmp/pagecount_mrjob.hadoop.20250527.094917.716969/output...
"p2"      2
"p4"      2
"p3"      2
"p1"      4
"p6"      1
Removing temp directory /tmp/pagecount_mrjob.hadoop.20250527.094917.716969...

```

```

GNU nano 6.2
from mrjob.job import MRJob

class MRPageCount(MRJob):

    def mapper(self, _, line):

        line = line.strip()

        words = line.split()
        for word in words:
            yield (word.lower(), 1)

    def reducer(self, word, counts):

        yield (word, sum(counts))

if __name__ == '__main__':
    MRPageCount.run()

```

OUTPUT:

```

Streaming final output from /tmp/pagecount_mrjob.hadoop.20250527.094917.716969/output...
"p2"      2
"p4"      2
"p3"      2
"p1"      4
"p6"      1
Removing temp directory /tmp/pagecount_mrjob.hadoop.20250527.094917.716969...

```

8b)

Code+Output:

Connecting to mongoDB:

```
bnscece@bnscece-HP-Elite-Tower-800-G9-Desktop-PC:~$ mongosh
Current Mongosh Log ID: 6835ca3484a52b8b2dc08e7d
Connecting to:      mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+2.2.1
Using MongoDB:      7.0.7
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/

-----
  The server generated these startup warnings when booting
  2025-05-27T19:49:20.245+05:30: Using the XFS filesystem is strongly recommended with the WiredTiger storage engine. See http://dochub.mongodb.org/c
  2025-05-27T19:49:22.391+05:30: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted
-----

test> use myDB
switched to db myDB
```

INSERTING DATA:

```

myDB> db.Student.drop()
true
myDB> db.createCollection('Student')
{ ok: 1 }
myDB> db.Student.insert({_id:1, StudName:'AKshay', Grade:'VII', Hobbies:'Chess', DOJ:'11-11-11'});
{ acknowledged: true, insertedIds: { '0': 1 } }
myDB> db.Student.find()
[
  {
    _id: 1,
    StudName: 'AKshay',
    Grade: 'VII',
    Hobbies: 'Chess',
    DOJ: '11-11-11'
  }
]
myDB> db.Student.insert({_id:2, StudName:'ABC', Grade:'VI', Hobbies:'Games', DOJ:'12-11-11'});
{ acknowledged: true, insertedIds: { '0': 2 } }
myDB> db.Student.insert({_id:3, StudName:'ASD', Grade:'V', Hobbies:'Games', DOJ:'10-11-11'});
{ acknowledged: true, insertedIds: { '0': 3 } }
myDB> db.Student.insert({_id:4, StudName:'ASQWE', Grade:'VIII', Hobbies:'Games', DOJ:'10-11-11'});
{ acknowledged: true, insertedIds: { '0': 4 } }
myDB> db.Student.insert({_id:5, StudName:'ABCD5', Grade:'VII', Hobbies:'Games', DOJ:'10-11-11'});
{ acknowledged: true, insertedIds: { '0': 5 } }
myDB> db.Student.find()
\[
  {
    _id: 1,
    StudName: 'AKshay',
    Grade: 'VII',
    Hobbies: 'Chess',
    DOJ: '11-11-11'
  },
  {
    _id: 2,
    StudName: 'ABC',
    Grade: 'VI',
    Hobbies: 'Games',
    DOJ: '12-11-11'
  },
  {
    _id: 3,
    StudName: 'ASD',
    Grade: 'V',
    Hobbies: 'Games',
    DOJ: '10-11-11'
  },
  {
    _id: 4,
    StudName: 'ASQWE',
    Grade: 'VIII',
    Hobbies: 'Games',
    DOJ: '10-11-11'
  },
  {
    _id: 5,
    StudName: 'ABCD5',
    Grade: 'VII',
    Hobbies: 'Games',
    DOJ: '10-11-11'
  }
]

```

1)

```

myDB> db.Student.find({}, {StudName:1});
[
  { _id: 1, StudName: 'AKshay' },
  { _id: 2, StudName: 'ABC' },
  { _id: 3, StudName: 'ASD' },
  { _id: 4, StudName: 'ASQWE' },
  { _id: 5, StudName: 'ABCD5' }
]
myDB> db.Student.find({}, {StudName:1, _id:0});
[
  { StudName: 'AKshay' },
  { StudName: 'ABC' },
  { StudName: 'ASD' },
  { StudName: 'ASQWE' },
  { StudName: 'ABCD5' }
]

```

2)

```

myDB> db.Student.find({_id:1}, {StudName:1, Grade:1, _id:1});
\[ { _id: 1, StudName: 'AKshay', Grade: 'VII' } ]
myDB>

```

3)NEXT PAGE

```
myDB> db.Student.find({Grade:{$ne:'VIII'}});
[
  {
    _id: 1,
    StudName: 'AKshay',
    Grade: 'VII',
    Hobbies: 'Chess',
    DOJ: '11-11-11'
  },
  {
    _id: 2,
    StudName: 'ABC',
    Grade: 'VI',
    Hobbies: 'Games',
    DOJ: '12-11-11'
  },
  {
    _id: 3,
    StudName: 'ASD',
    Grade: 'V',
    Hobbies: 'Games',
    DOJ: '10-11-11'
  },
  {
    _id: 5,
    StudName: 'ABCDs',
    Grade: 'VIII'
```

```
    Grade: 'VII',
    Hobbies: 'Chess',
    DOJ: '11-11-11'
  },
  {
    _id: 2,
    StudName: 'ABC',
    Grade: 'VI',
    Hobbies: 'Games',
    DOJ: '12-11-11'
  },
  {
    _id: 3,
    StudName: 'ASD',
    Grade: 'V',
    Hobbies: 'Games',
    DOJ: '10-11-11'
  },
  {
    _id: 5,
    StudName: 'ABCDs',
    Grade: 'VIII',
    Hobbies: 'Games',
    DOJ: '10-11-11'
  }
]
```

4)

```
myDB> db.Student.find({Hobbies:'Chess', StudName:'AKshay'});
[
  {
    _id: 1,
    StudName: 'AKshay',
    Grade: 'VII',
    Hobbies: 'Chess',
    DOJ: '11-11-11'
  }
]
```

5)

```
myDB> db.Student.find({StudName:/SS/});
[
  {
    _id: 5,
    StudName: 'ABCDS',
    Grade: 'VII',
    Hobbies: 'Games',
    DOJ: '10-11-11'
  }
]
myDB>
```