

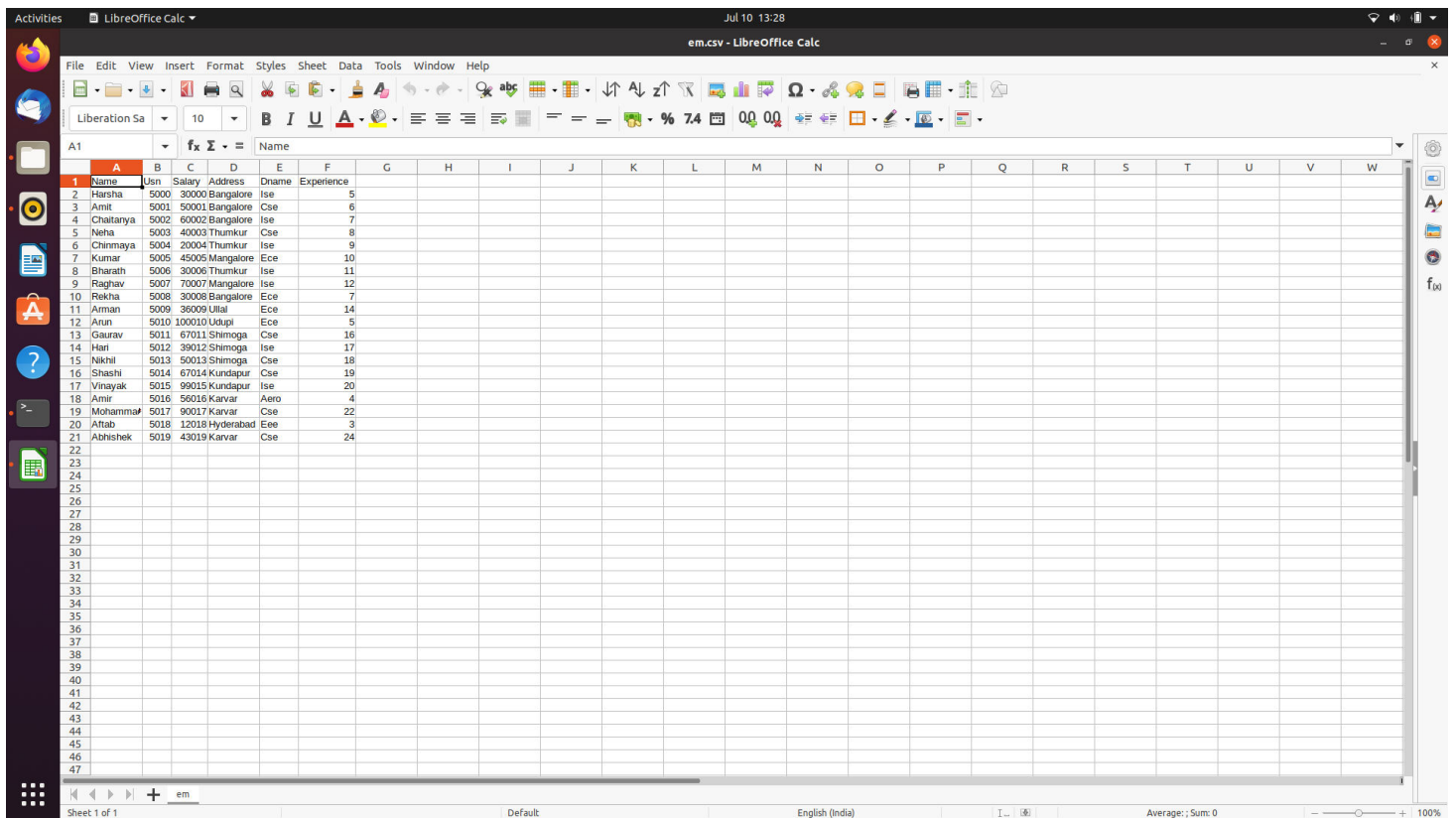
# LA2

Chinmayakrishna B

1NT18IS047

Exercise1: Hadoop map reduce

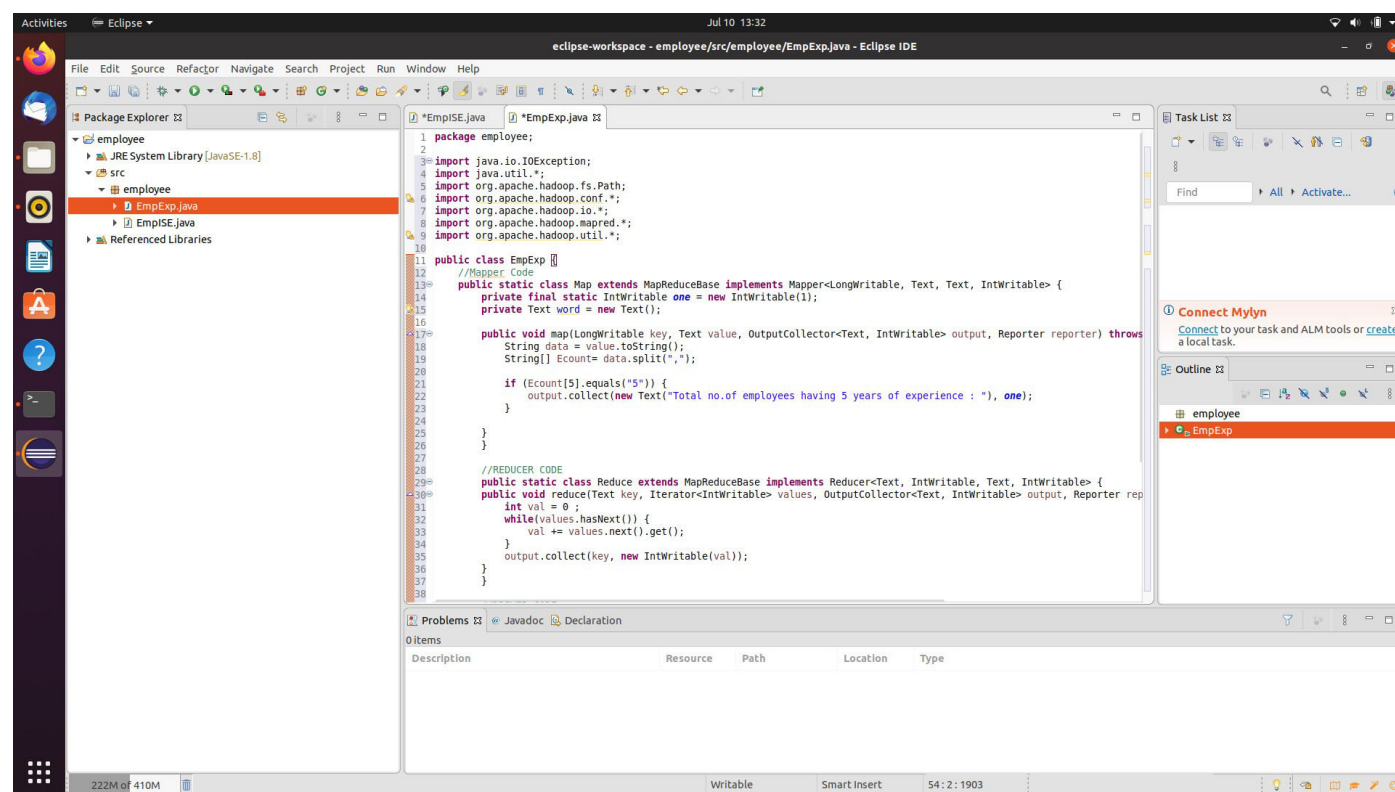
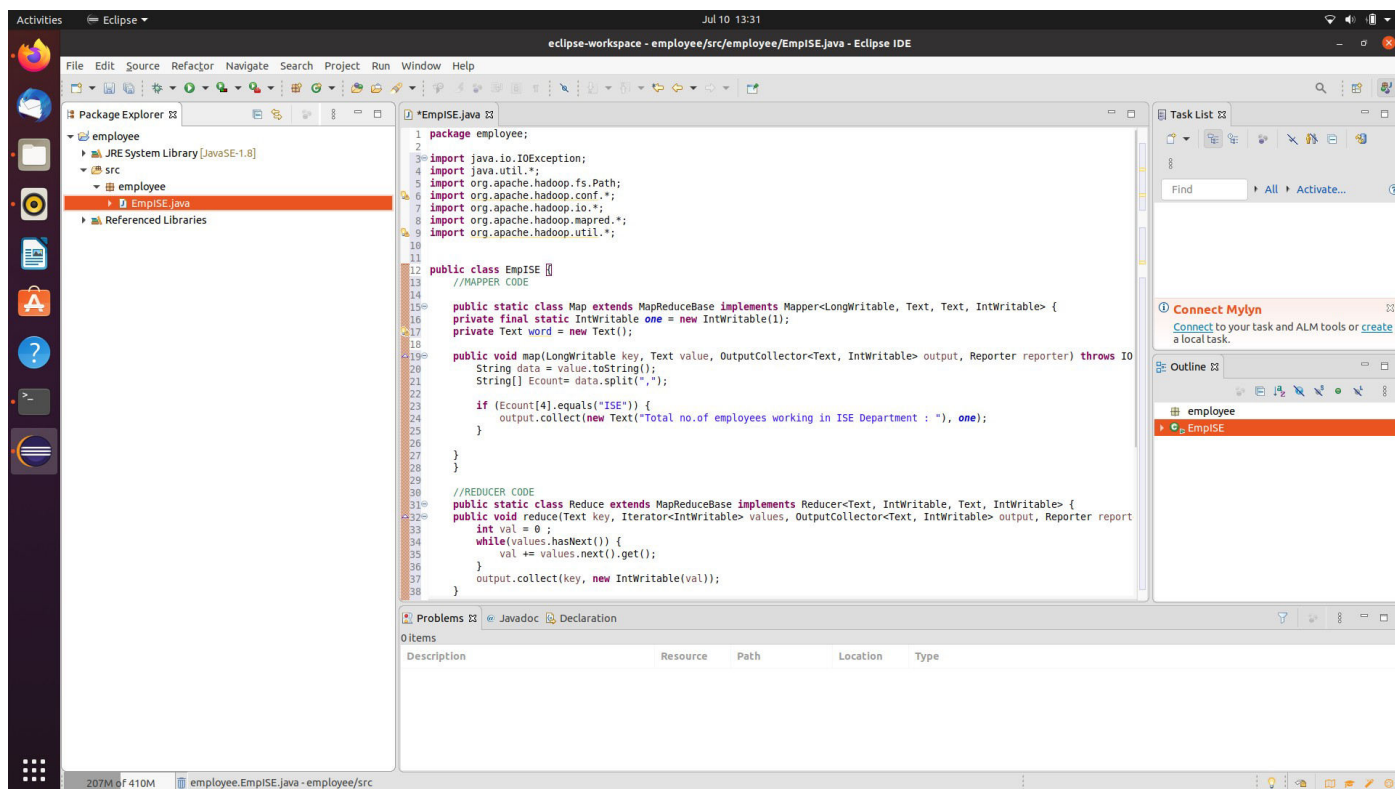
-> CSV file

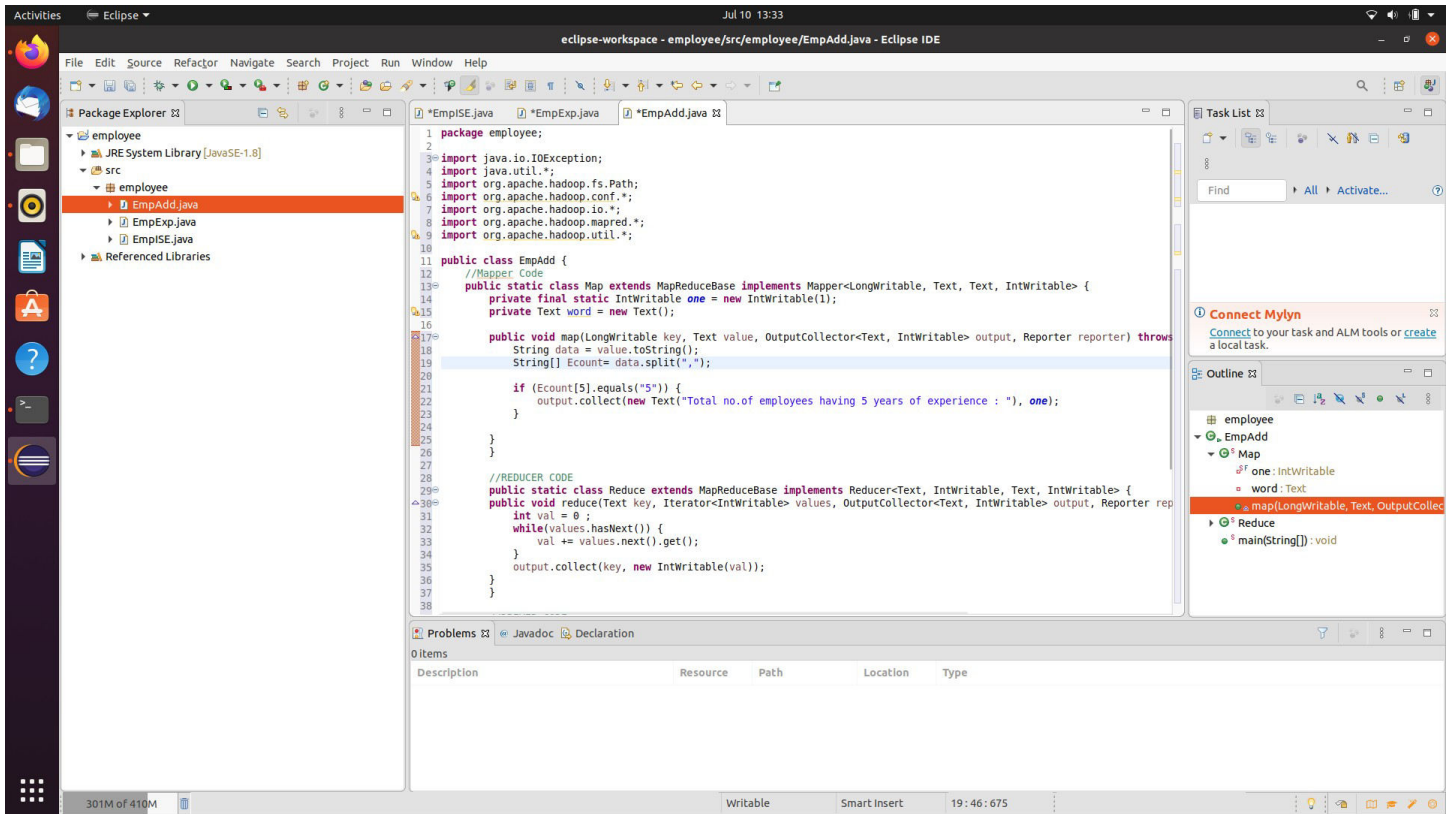


The screenshot shows the LibreOffice Calc application window with a CSV file named 'em.csv' open. The data is organized into columns: A (Name), B (Id), C (Salary), D (Address), E (Dname), and F (Experience). The data rows are numbered 1 through 21. The status bar at the bottom indicates 'Sheet 1 of 1', 'Default', 'English (India)', and 'Average: ; Sum: 0'.

	A	B	C	D	E	F
	Name	Id	Salary	Address	Dname	Experience
1	Harsha	5000	30000	Bangalore	Ise	5
2	Amit	5001	50000	Bangalore	Cse	6
3	Chaitanya	5002	60000	Bangalore	Ise	7
4	Neha	5003	40000	Thumkur	Cse	8
5	Chinmaya	5004	20004	Thumkur	Ise	9
6	Kumar	5005	45005	Mangalore	Ece	10
7	Bharath	5006	30006	Thumkur	Ise	11
8	Raghav	5007	70007	Mangalore	Ise	12
9	Rakha	5008	30008	Bangalore	Ece	7
10	Arman	5009	36009	Uttal	Ece	14
11	Arun	5010	100010	Udupi	Ece	5
12	Gaurav	5011	67011	Shimoga	Cse	16
13	Hari	5012	39012	Shimoga	Ise	17
14	Nikhil	5013	50013	Shimoga	Cse	18
15	Shashi	5014	67014	Kundapur	Cse	19
16	Vinayak	5015	99015	Kundapur	Ise	20
17	Amir	5016	56016	Karvar	Aero	4
18	Mohammad	5017	90017	Karvar	Cse	22
19	Aftab	5018	12018	Hyderabad	Ece	3
20	Abhishek	5019	43019	Karvar	Cse	24

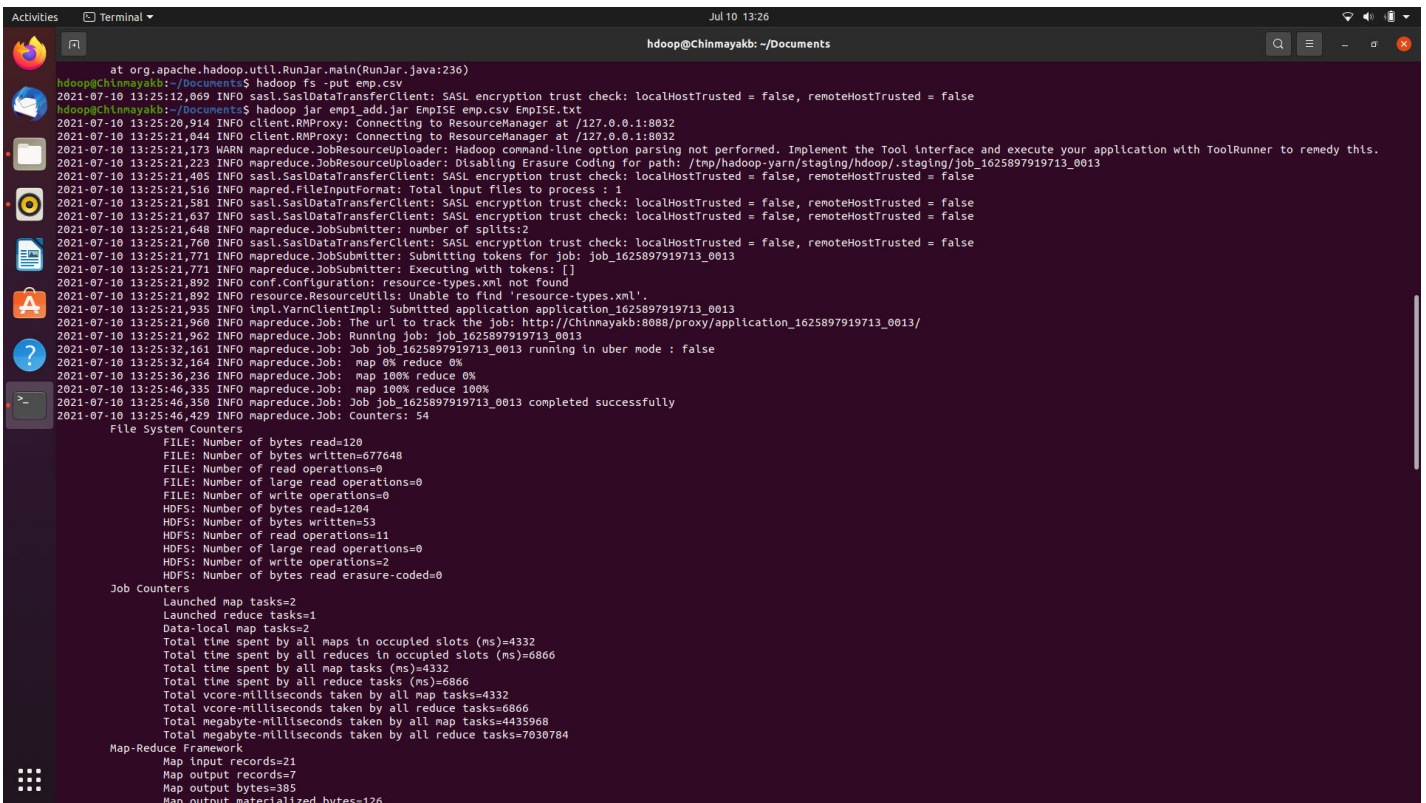
-> Eclipse





## Task

1) Total number of employees who work in ISE department





```
Activities Terminal Jul 10 13:27 hdoop@Chinmayakb: ~/Documents

[-count [-q] [-h] [-v] [-t {<storage type>}] [-u] [-x] [-e] <path> ...]
[-cp [-f] [-p] [-ptopack]] [-d] <src> ... <dst>]
[-createSnapshot <snapshotDir> [-snapshotName]]
[-deleteSnapshot <snapshotDir> <snapshotName>]
[-df [-h] <path> ...]
[-du [-s] [-h] [-v] [-x] <path> ...]
[-expunge [-immediate]]
[-find <path> ... <expressions> ...]
[-get [-f] [-p] [-ignoreCrc] [-crc] <src> ... <localdst>]
[-getfacl [-R] <path>]
[-getfattr [-R] [-n name | -d] [-e en] <path>]
[-getmerge [-nl] [-skip-empty-file] <src> <localdst>]
[-head <file>]
[-help [cmd ...]]
[-ls [-C] [-d] [-h] [-q] [-R] [-t] [-S] [-r] [-u] [-e] <path> ...]
[-mkdir [-p] <path> ...]
[-moveFromLocal <localSrc> ... <dst>]
[-moveToLocal <src> <localdst>]
[-mv <src> ... <dst>]
[-put [-f] [-p] [-t]] [-d] <localSrc> ... <dst>]
[-renameSnapshot <snapshotDir> <oldName> <newName>]
[-rm [-f] [-r] [-R] [-skipTrash] [-safely] <src> ...]
[-rmdir [--ignore-fail-on-non-empty] <dir> ...]
[-setfacl [-R] [--bl-k] [--w-x <acl_spec>] <path>][--set <acl_spec> <path>]]
[-setfattr [-n name [-v value] | -t <x_name>] <path>]
[-setrep [-R] [-w] <rep> <path> ...]
[-stat [format] <path> ...]
[-tail [-f] [-s <sleep interval>] <file>]
[-test [-defswrz] <path>]
[-text [-ignoreCrc] <src> ...]
[-touch [-a] [-m] [-t TIMESTAMP] [-c] <path> ...]
[-touchz <path> ...]
[-truncate [-w] <length> <path> ...]
[-usage [cmd ...]]

Generic options supported are:
-conf <configuration file> specify an application configuration file
-D <property=value> define a value for a given property
-fs <file:///hdfs://namenode:port> specify default filesystem URL to use, overrides 'fs.defaultFS' property from configurations.
-jt <local|resource|manager:port> specify a ResourceManager
-files <file1,...> specify a comma-separated list of files to be copied to the map reduce cluster
-libjars <jar1,...> specify a comma-separated list of jar files to be included in the classpath
-archives <archive1,...> specify a comma-separated list of archives to be unarchived on the compute machines

The general command line syntax is:
command [genericOptions] [commandOptions]

hdoop@Chinmayakb: ~/Documents$ hadoop fs -ls EmpISE.txt
Found 2 items
-rw-r--r-- 1 hdoop supergroup 0 2021-07-10 13:25 EmpISE.txt/_SUCCESS
-rw-r--r-- 1 hdoop supergroup 53 2021-07-10 13:25 EmpISE.txt/part-00000

hdoop@Chinmayakb: ~/Documents$ hadoop fs -cat EmpISE.txt/part-00000
2021-07-10 13:25:25,259 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
Total no. of employees working in ISE Department : 7
hdoop@Chinmayakb: ~/Documents$
```

2)Total Number of employees with experience=5years

```
Activities Terminal Jul 10 13:36 hdoop@Chinmayakb: ~/Documents

drwxr-xr-x - hdoop supergroup 0 2021-07-10 13:25 EmpISE.txt
-rw-r--r-- 1 hdoop supergroup 680 2021-07-10 13:25 emp.csv
hdoop@Chinmayakb: ~/Documents$ ls
1NT181S047_Chinmayakrishna_BD_LA2-ex2.odt em.csv emp1.jar emp.csv EmpExp.jar
hdoop@Chinmayakb: ~/Documents$ hadoop jar empExp.jar EmpExp emp.csv EmpExp
JAR does not exist or is not a normal file: /home/hdoop/Documents/empExp.jar
hdoop@Chinmayakb: ~/Documents$ hadoop jar empExp.jar EmpExp emp.csv EmpExp
2021-07-10 13:35:44,107 INFO client.RMProxy: Connecting to ResourceManager at /127.0.0.1:8032
2021-07-10 13:35:44,245 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-07-10 13:35:44,377 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
2021-07-10 13:35:44,469 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/hdoop/.staging/job_1625897919713_0014
2021-07-10 13:35:44,659 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-07-10 13:35:44,770 INFO mapred.FileInputFormat: Total input files to process : 1
2021-07-10 13:35:44,836 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-07-10 13:35:44,892 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-07-10 13:35:44,903 INFO mapreduce.JobSubmitter: number of splits:2
2021-07-10 13:35:45,148 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-07-10 13:35:45,159 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1625897919713_0014
2021-07-10 13:35:45,159 INFO mapreduce.JobSubmitter: Executing with tokens: []
2021-07-10 13:35:45,305 INFO conf.Configuration: resource-types.xml not found
2021-07-10 13:35:45,349 INFO impl.YarnClientImpl: Submitted application application_1625897919713_0014
2021-07-10 13:35:45,378 INFO mapreduce.Job: The url to track the job: http://Chinmayakb:8088/proxy/application_1625897919713_0014/
2021-07-10 13:35:45,379 INFO mapreduce.Job: Running job: job_1625897919713_0014
2021-07-10 13:35:50,512 INFO mapreduce.Job: Job job_1625897919713_0014 running in uber mode : false
2021-07-10 13:35:50,515 INFO mapreduce.Job: map 0% reduce 0%
2021-07-10 13:35:54,579 INFO mapreduce.Job: map 100% reduce 0%
2021-07-10 13:35:58,624 INFO mapreduce.Job: map 100% reduce 100%
2021-07-10 13:35:59,656 INFO mapreduce.Job: Job job_1625897919713_0014 completed successfully
2021-07-10 13:35:59,744 INFO mapreduce.Job: Counters: 54

File System Counters
  FILE: Number of bytes read=126
  FILE: Number of bytes written=677657
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=1204
  HDFS: Number of bytes written=56
  HDFS: Number of read operations=11
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=2
  HDFS: Number of bytes read erasure-coded=0

Job Counters
  Launched map tasks=2
  Launched reduce tasks=1
  Data-local map tasks=2
  Total time spent by all maps in occupied slots (ms)=3918
  Total time spent by all reduces in occupied slots (ms)=1705
  Total time spent by all map tasks (ms)=3918
  Total time spent by all reduce tasks (ms)=1705
  Total vcore-milliseconds taken by all map tasks=3918
  Total vcore-milliseconds taken by all reduce tasks=1705
  Total megabyte-milliseconds taken by all map tasks=4012032
  Total megabyte-milliseconds taken by all reduce tasks=1745920

Map-Reduce Framework
  Map input records=21
```

```
Activities Terminal Jul 10 13:36 hdoop@Chinnmayakb: ~/Documents

Total time spent by all map tasks (ms)=3918
Total time spent by all reduce tasks (ms)=1705
Total vcore-milliseconds taken by all map tasks=3918
Total vcore-milliseconds taken by all reduce tasks=1705
Total megabyte-milliseconds taken by all map tasks=4012032
Total megabyte-milliseconds taken by all reduce tasks=1745920

Map-Reduce Framework
  Map input records=21
  Map output records=8
  Map output bytes=464
  Map output materialized bytes=132
  Input split bytes=184
  Combine input records=8
  Combine output records=2
  Reduce input groups=1
  Reduce shuffle bytes=132
  Reduce input records=2
  Reduce output records=1
  Spilled Records=4
  Shuffled Maps =2
  Failed Shuffles=0
  Merged Map outputs=2
  GC time elapsed (ms)=120
  CPU time spent (ms)=1240
  Physical memory (bytes) snapshot=904175616
  Virtual memory (bytes) snapshot=7668494336
  Total committed heap usage (bytes)=770179072
  Peak Map Physical memory (bytes)=237760256
  Peak Map Virtual memory (bytes)=2553155584
  Peak Reduce Physical memory (bytes)=230051840
  Peak Reduce Virtual memory (bytes)=2562248704

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=1020
File Output Format Counters
  Bytes Written=56

hdoop@Chinnmayakb:~/Documents$ hadoop fs -ls
Found 3 items
drwxr-xr-x - hdoop supergroup 0 2021-07-10 13:35 EmpExp
drwxr-xr-x - hdoop supergroup 0 2021-07-10 13:25 EmpISE.txt
-rw-r--r-- 1 hdoop supergroup 680 2021-07-10 13:25 emp.csv
hdoop@Chinnmayakb:~/Documents$ hadoop fs -ls EmpExp
Found 2 items
-rw-r--r-- 1 hdoop supergroup 0 2021-07-10 13:35 EmpExp_SUCCESS
-rw-r--r-- 1 hdoop supergroup 56 2021-07-10 13:35 EmpExp/part-00000
hdoop@Chinnmayakb:~/Documents$ hadoop fs -cat EmpExp/part-00000
2021-07-10 13:36:31,453 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
Total no. of employees having 2 years of experience : 8
hdoop@Chinnmayakb:~/Documents$
```

### 3) Count the number of employees who lives in Bangalore

```
Activities Terminal Jul 10 13:39 hdoop@Chinnmayakb: ~/Documents

at java.lang.ClassLoader.loadClass(ClassLoader.java:351)
at java.lang.Class.forName0(Native Method)
at java.lang.Class.forName(Class.java:348)
at org.apache.hadoop.util.RunJar.run(RunJar.java:316)
at org.apache.hadoop.util.RunJar.main(RunJar.java:236)
hdoop@Chinnmayakb:~/Documents$ hadoop jar EmpAdd.jar ExpAdd emp.csv EmpAdd
2021-07-10 13:38:32,107 INFO client.RMProxy: Connecting to ResourceManager at /127.0.0.1:8032
2021-07-10 13:38:32,234 INFO client.RMProxy: Connecting to ResourceManager at /127.0.0.1:8032
2021-07-10 13:38:32,369 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
2021-07-10 13:38:32,411 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/hdoop/staging/job_1625897919713_0015
2021-07-10 13:38:32,495 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-07-10 13:38:32,592 INFO mapred.FileInputFormat: Total input files to process : 1
2021-07-10 13:38:32,656 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-07-10 13:38:32,768 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-07-10 13:38:32,779 INFO mapreduce.JobSubmitter: number of splits:2
2021-07-10 13:38:32,924 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-07-10 13:38:32,934 INFO mapreduce.JobSubmitter: Submitting tokens for Job: job_1625897919713_0015
2021-07-10 13:38:32,935 INFO mapreduce.JobSubmitter: Executing with tokens: []
2021-07-10 13:38:33,055 INFO conf.Configuration: resource-types.xml not found
2021-07-10 13:38:33,055 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2021-07-10 13:38:33,097 INFO impl.YarnClientImpl: Submitted application application_1625897919713_0015
2021-07-10 13:38:33,122 INFO mapreduce.Job: The url to track the job: https://chinnmayakb:8088/proxy/application_1625897919713_0015/
2021-07-10 13:38:33,123 INFO mapreduce.Job: Running job: job_1625897919713_0015
2021-07-10 13:38:38,198 INFO mapreduce.Job: Job job_1625897919713_0015 running in uber mode : false
2021-07-10 13:38:38,200 INFO mapreduce.Job: map 0% reduce 0%
2021-07-10 13:38:42,254 INFO mapreduce.Job: map 100% reduce 0%
2021-07-10 13:38:46,383 INFO mapreduce.Job: map 100% reduce 100%
2021-07-10 13:38:47,336 INFO mapreduce.Job: Job job_1625897919713_0015 completed successfully
2021-07-10 13:38:47,425 INFO mapreduce.Job: Counters: 54

File System Counters
  FILE: Number of bytes read=58
  FILE: Number of bytes written=677497
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=1204
  HDFS: Number of bytes written=48
  HDFS: Number of read operations=11
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=2
  HDFS: Number of bytes read erasure-coded=0

Job Counters
  Launched map tasks=2
  Launched reduce tasks=1
  Data-local map tasks=2
  Total time spent by all maps in occupied slots (ms)=4046
  Total time spent by all reduces in occupied slots (ms)=1686
  Total time spent by all map tasks (ms)=4046
  Total time spent by all reduce tasks (ms)=1686
  Total vcore-milliseconds taken by all map tasks=4046
  Total vcore-milliseconds taken by all reduce tasks=1686
  Total megabyte-milliseconds taken by all map tasks=4143104
  Total megabyte-milliseconds taken by all reduce tasks=1726464

Map-Reduce Framework
  Map input records=21
  Map output records=1
```

```

Total time spent by all reduce tasks (ms)=1686
Total vcore-millisecods taken by all map tasks=4046
Total vcore-millisecods taken by all reduce tasks=1686
Total megabyte-millisecods taken by all map tasks=4143104
Total megabyte-millisecods taken by all reduce tasks=1726464
Map-Reduce Framework
  Map input records=21
  Map output records=1
  Map output bytes=50
  Map output materialized bytes=64
  Input split bytes=184
  Combine input records=1
  Combine output records=1
  Reduce input groups=1
  Reduce shuffle bytes=64
  Reduce input records=1
  Reduce output records=1
  Spilled Records=2
  Shuffled Maps =2
  Failed Shuffles=0
  Merged Map outputs=2
  GC time elapsed (ms)=121
  CPU time spent (ms)=1200
  Physical memory (bytes) snapshot=905273344
  Virtual memory (bytes) snapshot=7607679232
  Total committed heap usage (bytes)=772276224
  Peak Map Physical memory (bytes)=336797696
  Peak Map Virtual memory (bytes)=2552590336
  Peak Reduce Physical memory (bytes)=231878656
  Peak Reduce Virtual memory (bytes)=2562584576
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=1020
  File Output Format Counters
    Bytes Written=48
hadoop@Chinnayakb:~/Documents$ hadoop fs -ls
Found 4 items
drwxr-xr-x  - hadoop supergroup      0 2021-07-10 13:38 EmpAdd
drwxr-xr-x  - hadoop supergroup      0 2021-07-10 13:35 EmpExp
drwxr-xr-x  - hadoop supergroup      0 2021-07-10 13:25 EmpISE.txt
-rw-r--r--  1 hadoop supergroup    680 2021-07-10 13:25 emp.csv
hadoop@Chinnayakb:~/Documents$ hadoop fs -ls EmpAdd
Found 2 items
-rw-r--r--  1 hadoop supergroup      0 2021-07-10 13:38 EmpAdd/_SUCCESS
-rw-r--r--  1 hadoop supergroup    48 2021-07-10 13:38 EmpAdd/part-00000
hadoop@Chinnayakb:~/Documents$ hadoop fs -cat EmpAdd/part-00000
2021-07-10 13:39:15,595 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
Total no.of employees working in Bengaluru : 1
hadoop@Chinnayakb:~/Documents$

```

## Task1:Code:

```

package employee;
import java.io.IOException;
import java.util.*;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.conf.*;
import org.apache.hadoop.io.*;
import org.apache.hadoop.mapred.*;
import org.apache.hadoop.util.*;

```

```

public class EmpISE {
    //MAPPER CODE

```

```

    public static class Map extends MapReduceBase implements Mapper<LongWritable, Text,
Text, IntWritable> {
        private final static IntWritable one = new IntWritable(1);
        private Text word = new Text();

```

```

        public void map(LongWritable key, Text value, OutputCollector<Text, IntWritable> output,
Reporter reporter) throws IOException {
            String data = value.toString();
            String[] Ecount= data.split(",");

            if (Ecount[4].equals("ISE")) {
                output.collect(new Text("Total no.of employees working in ISE Department :
"), one);

```

```

    }

}
}

//REDUCER CODE
public static class Reduce extends MapReduceBase implements Reducer<Text, IntWritable,
Text, IntWritable> {
    public void reduce(Text key, Iterator<IntWritable> values, OutputCollector<Text,
IntWritable> output, Reporter reporter) throws IOException { //{little: {1,1}}
        int val = 0 ;
        while(values.hasNext()) {
            val += values.next().get();
        }
        output.collect(key, new IntWritable(val));
    }
}

//DRIVER CODE
public static void main(String[] args) throws Exception {
    JobConf conf = new JobConf(EmpISE.class);
    conf.setJobName("Total no.of employees working in ISE Department");
    conf.setOutputKeyClass(Text.class);
    conf.setOutputValueClass(IntWritable.class);
    conf.setMapperClass(Map.class);
    conf.setCombinerClass(Reduce.class);
    conf.setReducerClass(Reduce.class);
    conf.setInputFormat(TextInputFormat.class);
    conf.setOutputFormat(TextOutputFormat.class);
    FileInputFormat.setInputPaths(conf, new Path(args[0]));
    FileOutputFormat.setOutputPath(conf, new Path(args[1]));
    JobClient.runJob(conf);
}
}

```

## Task2:code:

```

package employee;

import java.io.IOException;
import java.util.*;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.conf.*;
import org.apache.hadoop.io.*;
import org.apache.hadoop.mapred.*;
import org.apache.hadoop.util.*;

public class EmpExp {
    //Mapper Code

```

```

    public static class Map extends MapReduceBase implements Mapper<LongWritable,
Text, Text, IntWritable> {
        private final static IntWritable one = new IntWritable(1);
        private Text word = new Text();

        public void map(LongWritable key, Text value, OutputCollector<Text,
IntWritable> output, Reporter reporter) throws IOException {
            String data = value.toString();
            String[] Ecount= data.split(",");

            if (Ecount[5].equals("5")) {
                output.collect(new Text("Total no.of employees having 5 years
of experience : "), one);
            }

        }
    }

```

```

//REDUCER CODE
    public static class Reduce extends MapReduceBase implements
Reducer<Text, IntWritable, Text, IntWritable> {
        public void reduce(Text key, Iterator<IntWritable> values,
OutputCollector<Text, IntWritable> output, Reporter reporter) throws IOException {
//{little: {1,1}}

            int val = 0 ;
            while(values.hasNext()) {
                val += values.next().get();
            }
            output.collect(key, new IntWritable(val));
        }
    }

```

```

//DRIVER CODE
    public static void main(String[] args) throws Exception {
        JobConf conf = new JobConf(EmpExp.class);
        conf.setJobName("Total no.of employees having 5 years of
experience");

        conf.setOutputKeyClass(Text.class);
        conf.setOutputValueClass(IntWritable.class);
        conf.setMapperClass(Map.class);
        conf.setCombinerClass(Reduce.class);
        conf.setReducerClass(Reduce.class);
        conf.setInputFormat(TextInputFormat.class);
        conf.setOutputFormat(TextOutputFormat.class);
        FileInputFormat.setInputPaths(conf, new Path(args[0]));
        FileOutputFormat.setOutputPath(conf, new Path(args[1]));
        JobClient.runJob(conf);
    }

```



```
}
```

### Task3:Code:

```
import java.io.IOException;
import java.util.*;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.conf.*;
import org.apache.hadoop.io.*;
import org.apache.hadoop.mapred.*;
import org.apache.hadoop.util.*;

public class ExpAdd {
    //Mapper Code
    public static class Map extends MapReduceBase implements Mapper<LongWritable,
Text, Text, IntWritable> {
        private final static IntWritable one = new IntWritable(1);
        private Text word = new Text();

        public void map(LongWritable key, Text value, OutputCollector<Text,
IntWritable> output, Reporter reporter) throws IOException {
            String data = value.toString();
            String[] Ecount= data.split(",");

            if (Ecount[3].equals("Bengaluru")) {
                output.collect(new Text("Total no.of employees working in
Bengaluru : "), one);
            }

        }
    }

    //REDUCER CODE
    public static class Reduce extends MapReduceBase implements
Reducer<Text, IntWritable, Text, IntWritable> {
        public void reduce(Text key, Iterator<IntWritable> values,
OutputCollector<Text, IntWritable> output, Reporter reporter) throws IOException {
            //{little: {1,1}}

            int val = 0 ;
            while(values.hasNext()) {
                val += values.next().get();
            }
            output.collect(key, new IntWritable(val));
        }
    }

    //DRIVER CODE
```

```
public static void main(String[] args) throws Exception {  
    JobConf conf = new JobConf(EmpExp.class);  
    conf.setJobName("Total no.of employees working in Bengaluru");  
    conf.setOutputKeyClass(Text.class);  
    conf.setOutputValueClass(IntWritable.class);  
    conf.setMapperClass(Map.class);  
    conf.setCombinerClass(Reduce.class);  
    conf.setReducerClass(Reduce.class);  
    conf.setInputFormat(TextInputFormat.class);  
    conf.setOutputFormat(TextOutputFormat.class);  
    FileInputFormat.setInputPaths(conf, new Path(args[0]));  
    FileOutputFormat.setOutputPath(conf, new Path(args[1]));  
    JobClient.runJob(conf);  
}
```

```
}
```